

Dr. Barbara Höhle (2002)

Der Einstieg in die Grammatik: Die Rolle der Phonologie/Syntax-Schnittstelle für Sprachverarbeitung und Spracherwerb

Habilitationsschrift. Freie Universität Berlin

Alle Rechte vorbehalten, all rights reserved, 2002 Dr. Barbara Höhle.

Dieses Dokument einschließlich aller seiner Teile ist urheberrechtlich geschützt. Bei Zitaten ist die Autorin zu benennen. Für fehlerhafte Angaben und deren Folgen übernimmt die Autorin weder eine juristische Verantwortung noch irgendeine Haftung.

0	Einleitung	1
1	Grundlagen des Spracherwerbs	7
1.1	Der spezifische nativistische Ansatz.....	9
1.2	Der generelle Nativismus.....	14
1.3	Spracherwerb und Sprachverarbeitung	18
1.3.1	Das System der Operating Principles	19
1.3.2	Bootstrapping-Mechanismen.....	20
1.4	Zwei grundlegende Verarbeitungsschritte im frühen Spracherwerb: Segmentierung und syntaktische Kategorisierung.....	23
1.4.1	Segmentierung des akustischen Signals	24
1.4.2	Syntaktische Kategorisierung sprachlicher Einheiten	26
2	Methoden zur Untersuchung früher sprachlicher und kognitiver Fähigkeiten	33
2.1	Habituationparadigmen	35
2.1.1	Das High-Amplitude Sucking Paradigma	35
2.1.2	Visuelle Fixation.....	38
2.2	Konditionierungsparadigmen	39
2.2.1	Das Conditioned Headturn Verfahren.....	39
2.3	Präferenzparadigmen	42
2.3.1	Das Headturn Preference Paradigma	42
2.3.2	Preferential Looking.....	44
2.4	Familiaritäts- oder Neuigkeitseffekt: Die Richtung von Präferenzen	46
3	Prosodische Information und Sprachverarbeitung	53
3.1	Lexikalische Verarbeitung.....	54
3.1.1	Rhythmische Segmentierung als Teil der lexikalischen Verarbeitung.....	54
3.1.2	Inputinformation für die rhythmische Segmentierung	60
3.1.2.1	Phonetische Korrelate der Wortbetonung.....	60
3.1.2.2	Phonetische Korrelate von Rhythmustypen	61
3.2	Syntaktische Verarbeitung	65
3.2.1	Prosodische Hinweise auf grammatische Klassen.....	65
3.2.2	Prosodische Phrasierung als Hinweis auf die syntaktische Struktur.....	68
3.2.3	Inputinformation für die Erkennung prosodischer Grenzen	74
4	Der prosodische Einstieg in den Spracherwerb	77
4.1	Sensitivität für prosodische Information	78

4.1.1	Sensitivität für globale prosodische Information	78
4.1.2	Sensitivität für spezifische prosodische Merkmale	83
4.1.2.1	Silbenzahl.....	83
4.1.2.2	Betonungsmuster.....	84
4.1.2.3	Grenzsignale.....	85
4.2	Die Ausnutzung prosodischer Merkmale für die Gliederung des Signals	88
4.2.1	Verarbeitung prosodischer Information und Erwerb lexikalischer Information.....	88
4.2.1.1	Wortsegmentierung.....	88
4.2.1.2	Experiment 1: Detektion trochäischer Zweisilbler bei deutschen Säuglingen.....	92
4.2.1.3	Experiment 2: Detektion trochäischer Zweisilbler in deutschen Stimuli durch englische Säuglinge.....	96
4.2.2	Die Trochäuspräferenz: Resultat einer Anpassung an die Umgebungssprache?.....	98
4.2.2.1	Experiment 3: Trochäuspräferenz bei 6 Monate alten Säuglingen.....	102
4.2.2.2	Experiment 4: Trochäuspräferenz bei 3 Monate alten Säuglingen.....	104
4.2.3	Prosodische Information und Wortkategorisierung.....	107
4.2.4	Prosodische Information und der Erwerb syntaktischen Wissens.....	108
4.2.4.1	Prosodische Information und Phrasensegmentierung.....	108
4.2.4.2	Prosodische Information und der Erwerb von Wortstellungsregeln	110
4.3	Möglichkeiten und Grenzen des Prosodischen Bootstrappings.....	115
4.3.1	Können akustische prosodische Markierungen rein datengetrieben verarbeitet werden?.....	115
4.3.1.1	Multifunktionalität akustisch-prosodischer Merkmale	115
4.3.1.2	Sprachspezifische Aspekte prosodischer Markierungen und prosodischer Struktur.....	117
4.3.2	Reliabilität prosodischer Merkmale als Grenzmarkierung	120
4.3.2.1	Die Zuverlässigkeit akustischer Grenzmarkierungen.....	120
4.3.2.2	Die Eigenständigkeit von prosodischer und syntaktischer Struktur.....	123
5	Segmentale und segmental-sequentielle Information als strukturelle Hinweise	127
5.1	Segmentale und sequentielle Information für strukturelle Grenzen	127
5.1.1	Akustische Hinweise.....	129
5.1.2	Übergangswahrscheinlichkeiten zwischen Phonemen und Silben.....	130
5.1.3	Phonotaktische Beschränkungen.....	136
5.1.4	Häufige Einheiten	138
5.1.5	Lexikalisches Wissen.....	142
5.2	Segmentale Information und Wortkategorisierung.....	145
5.3	Möglichkeiten und Grenzen der Nutzung segmental-sequentieller Information.....	148
6	Der segmental-distributionelle Erwerbsstrang	151
6.1	Voraussetzungen für den Erwerb segmental-distributionellen Wissens	152
6.1.1	Lautdiskrimination.....	152
6.1.2	Bildung von Repräsentationen.....	154
6.1.2.1	Experiment 5: Wiedererkennung unbetonter Elemente der geschlossenen Klasse im Kontext	157
6.1.3	Analysemechanismen für distributionelle Information.....	161
6.2	Der Aufbau sprachspezifischen distributionellen Wissens	164
6.2.1	Der Aufbau des Phoneminventars.....	164
6.2.2	Der Erwerb zielsprachlichen kombinatorischen und phonotaktischen Wissens.....	168
6.2.3	Der Aufbau erster Wortform-Repräsentationen	170
6.3	Die Nutzung segmental-distributionellen Wissens für die Verarbeitung des Inputs und den Aufbau höheren sprachlichen Wissens.....	171
6.3.1	Die Erkennung von Wortgrenzen.....	171
6.3.1.1	Allophonie	171
6.3.1.2	Statistisches Wissen über Lautabfolgen und phonotaktisches Wissen	174
6.3.1.3	Kookkurrenzen von Silben	176
6.3.1.4	Erkennung und Nutzung hochfrequenter Phonemfolgen	178
6.3.1.5	Nutzung von Funktionswörtern als Segmentierungshilfe	180
6.3.1.5.1	Experiment 6: Erkennung und Segmentierung von Artikeln bei 11- monatigen Kindern	181
6.3.1.5.2	Experiment 7: Segmentierung von Phrasen und jambischen Nomen bei 9 Monate alten Kindern.....	185

6.3.1.5.3	Experiment 8: Erkennung einer vertrauten Phonemsequenz oder eines Funktionswortes?...	188
6.3.1.6	Nutzung von Funktionswörtern für die syntaktische Klassifizierung.....	192
6.3.1.6.1	Experiment 9: Nutzung von Funktionswörtern zur syntaktischen Kategorisierung bei 15 Monate alten Kindern	194
6.3.1.6.2	Experiment 10: Nutzung von Funktionswörtern für die syntaktische Klassifizierung bei Kindern im Alter von 12 und 13 Monaten	199
6.4.	Der Beitrag segmental-sequentieller Information zur Lösung des Segmentierungs- und Klassifizierungsproblems	201
7	Neue Perspektiven auf den Spracherwerb.....	209
7.1	Früher Spracherwerb: Bestandsaufnahme und neue Perspektiven	210
7.1.1	Ein Modell des frühen Spracherwerbs	210
7.1.2	Die Perzeptions- und Analysemechanismen	218
7.1.2	Gewichtung und Integration von Hinweisreizen.....	223
7.1.3	Der Einfluss einzelsprachspezifischen Wissens auf die Sprachverarbeitung und den weiteren Spracherwerb	229
7.2.	Eine kritische Periode für den Spracherwerb?	233
Referenzen		237

0 Einleitung

The first twenty-five years of reserach on infant speech perception have provided many supprises, especially with respect to the skills that are present during the earliest stages of infancy and the rapidity with which experience with a particular language helps to shape and refine these skills. Given the range of issues that remain to be explored, the next twenty-five years promise to be just as exciting.

Peter W. Jusczyk, 1997, S. 231

In kaum einem Bereich hat sich unser Bild über die Fähigkeiten des Menschen in den letzten Jahrzehnten dramatischer verändert als für den Säugling. Während lange Zeit das Bild vorherrschte, der Säugling sei in seinen ersten Lebensmonaten ein eher passives Wesen, wissen wir heute, dass der Säugling quasi vom Tag seiner Geburt an – wenn nicht sogar schon pränatal – ein aktiver Lerner ist, der die Informationen seiner Umgebung verwertet, mit seiner Umgebung interagiert und bereits in seinen ersten Lebensmonaten über erstaunliches Wissen verfügt (als Überblicke s. Haith & Benson, 1998; Hennon, Hirsh-Pasek & Golinkoff, 2000). So besitzt bereits das Neugeborene effiziente Perzeptionsmechanismen für akustische Reize, die es ihm erlauben, Laute und eine Vielzahl anderer sprachlich relevanter Merkmale zu unterscheiden. Im ersten Lebensjahr verfügt der Säugling bereits über grundlegendes Wissen über Objekteigenschaften, er kann konzeptuelle Kategorien aufgrund gemeinsamer Merkmale von Objekten bilden, komplexe Handlungsabläufe in einzelne Komponenten gliedern und Ereignisse antizipieren. Er zeigt Verständnis für physikalische Grundprinzipien und arithmetische Operationen. Neben diesen kognitiven Fähigkeiten weist der Säugling vielfältige soziale Kompetenzen auf: er ist mit einer besonderen Aufmerksamkeit für sozial relevante Reize wie Gesichter ausgestattet und tritt schon während der ersten Lebensmonate mit seinen Bezugspersonen in Interaktion. In der zweiten Hälfte des ersten Lebensjahres reagiert der Säugling auch auf soziale Reize wie Zeigegesten und Blicke, folgt diesen und schafft somit ein gemeinsames

Aufmerksamkeitsfeld mit seinem Interaktionspartner. Wir wissen mittlerweile auch, dass sich im ersten Lebensjahr - also lange bevor das Kind seine ersten verstehbaren sprachlichen Äußerungen produziert - entscheidende Schritte für den Erwerb der Sprache vollziehen.

Zumindest für den Psycholinguisten stellt der Erwerb der Muttersprache den faszinierendsten Entwicklungsprozess dar, der sich während der ersten Lebensjahre vollzieht. Bei jedem Kind verblüfft aufs neue, wie selbstverständlich es unter normalen Bedingungen seine Muttersprache lernt. Die Verwunderung und Bewunderung liegt dabei vor allen Dingen bei denjenigen, die eine Vorstellung darüber haben, wie komplex und vielfältig die Regelgebilde und das Wissen sind, die die sprachliche Kompetenz ausmachen. Als kompetenter Muttersprachler muss der Sprachbenutzer Wissen über phonologische, syntaktische und morphologische Regularitäten der Sprache besitzen, er muss über ein mentales Lexikon verfügen, in dem beim erwachsenen Sprecher Wissen über bis zu 250.000 Wörter (Aitchinson, 1987), beim 6jährigen Kind bereits über ca. 15.000 Wörter vermutet wird (Clark, 1995). Für die in einer spezifischen Situation angemessene Sprachverwendung ist außerdem pragmatische Kompetenz, wie Wissen über Interaktionsmuster, Aspekte der Sprechsituation und der kommunikativen Intention des Sprechers und ihren Beziehungen zu den sprachlichen Ausdrucksmitteln, erforderlich.

So vielschichtig wie sich der Erwerbsgegenstand selbst darstellt, sind sicherlich die Faktoren, die einen erfolgreichen Verlauf des Spracherwerbs sicherstellen. Es ist kaum anzunehmen, dass der Erwerb eines derartig komplexen Systems wie der Sprache über einen einzigen Erwerbsmechanismus möglich ist, sondern dass eine Vielzahl von äußeren aber auch dem Lerner inhärenten Bedingungen erfüllt sein müssen. In der jüngeren Spracherwerbsforschung wird denn auch die Auffassung vertreten, dass der Säugling für den Aufbau seiner sprachlichen Kompetenz die unterschiedlichsten Quellen, die seine Umwelt ihm zur Verfügung stellt oder über die er bereits selbst verfügt, nutzt (Hirsh-Pasek & Golinkoff, 1996). Neben den sprachlichen Äußerungen, die der Säugling hört und verarbeitet und aus der er wichtige Informationen über die Merkmale seiner Sprache zieht, sind dies auch Aspekte der sozialen Interaktion (z. B. Snow, 1996; Tomasello, 1992). So ist inzwischen gut belegt, dass beispielsweise frühe Interaktionsmuster zwischen Mutter und Kind mit der Wortschatzentwicklung

korrelieren (z. B. Tomasello & Farrar, 1986). Ein erfolgreicher Spracherwerb ist zudem eingebunden in ein Netz verschiedener Komponenten der Kognition und der kognitiven Entwicklung, so beispielsweise in Fähigkeiten der Informationsverarbeitung und –speicherung (Gathercole & Baddeley, 1993) und den Aufbau konzeptuellen Wissens (s. die Beiträge in Bowerman & Levinson, 2001). Wir beginnen erst zu verstehen, welches die verschiedenen Bedingungsfaktoren sind, welchen Beitrag sie zu welchen Aspekten des Spracherwerbs leisten, in welcher Weise sie ineinandergreifen und in welchen Wechselbeziehungen sie zueinander stehen. Diese Fragen werden sicherlich im Zentrum der zukünftigen Spracherwerbsforschung stehen, denn erst ihre Beantwortung wird ein umfassendes Bild über die kindliche Fähigkeit zum Spracherwerb liefern.

Die vorliegende Arbeit beleuchtet einen Aspekt dieses komplexen Zusammenspiels externer und interner Bedingungsfaktoren für den Spracherwerb, speziell für den Erwerb phonologischen und syntaktischen Wissens. Es geht um die Frage, welche formalen Merkmale des sprachlichen Inputs vom Kind zu welchem Zeitpunkt der Entwicklung verarbeitet und für den Aufbau phonologischen und syntaktischen Wissens herangezogen werden können und es geht um eine genauere Beschreibung der hierfür genutzten Verarbeitungs- und Lernmechanismen. Dabei wird von den folgenden Grundannahmen ausgegangen:

- ?? Der sprachliche Input des Kindes enthält eine Vielzahl von Merkmalen, die den Erwerb dieses formalsprachlichen Wissens unterstützen.
- ?? Das Kind verfügt von Geburt an über spezifische Wahrnehmungs-, Verarbeitungs- und Lernmechanismen, die speziell die Ausnutzung dieser Informationstypen leisten.
- ?? Kinder und Erwachsene unterscheiden sich in ihren Sprachverarbeitungsmechanismen nicht prinzipiell; es ist aber ein zunehmender Einfluss des sprachlichen Wissens bei zunehmender sprachlicher Kompetenz auf die Verarbeitungsprozesse zu erwarten

Kapitel 1 gibt eine kurze Einführung in zwei konträre Positionen zu den Fähigkeiten, die das Kind in die Lösung der Erwerbsaufgabe einbringt. Dabei geht es in der jüngeren wissenschaftlichen Diskussion nicht mehr um die Frage, ob der Mensch überhaupt mit

spezifischen Voraussetzungen ausgestattet ist, die ihm den Spracherwerb ermöglichen, sondern es geht um die Frage, welcher Natur diese angeborenen Fähigkeiten sind, d. h. bestehen diese Voraussetzungen aus strukturellem oder prozeduralem Wissen, sind sie domänenspezifisch oder sind sie domänenübergreifend?

Unsere Betrachtungen konzentrieren sich auf die ersten eineinhalb Lebensjahre, da sich - wie erwähnt - bereits in dieser frühen Lebensphase wichtige Erwerbsschritte in Richtung auf das zielsprachliche phonologische und syntaktische System vollziehen. Die Untersuchungen zum frühen Spracherwerb erfordern spezifische experimentelle Techniken. Mittlerweile stehen eine Reihe von erprobten Untersuchungsverfahren zur Verfügung, in denen die Wahrnehmung und Verarbeitung von Sprache, die Anwendung bestimmter Lernmechanismen und die Entwicklung sprachlichen Wissens von Geburt an über die ersten beiden Lebensjahre untersucht werden kann. Kapitel 2 gibt einen Überblick über die gängigsten Paradigmen.

Kapitel 3 und 4 widmen sich mit ihrer Fokussierung auf prosodische Information einem Input-Parameter, dessen Relevanz für Sprachverarbeitung und Spracherwerb immer deutlicher wird. So sind Intonation, Rhythmus und Betonung für den Hörer wichtige Informationsquellen, mit deren Hilfe er strukturelle Einheiten im akustischen Signal identifizieren kann. So hilft der Rhythmus und die Betonung Wortgrenzen zu finden, Intonationsmerkmale wie Grundfrequenzverlauf und Dauerverhältnisse können für die Identifizierung von Satz- und Phrasengrenzen genutzt werden. Kapitel 3 gibt einen Überblick über den Forschungsstand zur Verarbeitung dieser Information bei Erwachsenen. Kapitel 4 fragt, ab wann der Säugling für diese Informationsquellen sensitiv ist und sie für den Spracherwerb nutzt. Es wird in diesem Zusammenhang auch über eigene Experimente berichtet, die zeigen, dass Säuglinge bereits im ersten Lebenshalbjahr das dominante Wortbetonungsmuster ihrer Muttersprache erkennen und dass sie dieses Wissen ohne weiteres spezifisches Wissen über andere Merkmale der Zielsprache für die Extraktion von Worteinheiten aus dem Input nutzen können.

Kapitel 5 und 6 konzentrieren sich auf die Nutzung und den Erwerb segmental-sequentiellen Wissens, in Form von beispielsweise phonotaktischen Regularitäten aber auch in Form von im Input häufigen Laut- oder Silbensequenzen, die als Einheit aus dem Input extrahiert und mental repräsentiert werden. Kapitel 5 befasst sich wiederum zunächst mit dem erwachsenen Sprachnutzer. Hier sehen wir, dass beispielsweise

phonotaktische Regularitäten und andere distributionelle Merkmale des Inputs vom Erwachsenen für die Sprachverarbeitung und auch den Spracherwerb genutzt werden. In diesem Zusammenhang wird auch die Rolle lexikalischen Wissens erörtert. Kapitel 6 widmet sich analogen Fähigkeiten beim Säugling. Hier wird vor allen Dingen der Frage nach dem Aufbau mentaler Repräsentationen für Funktionswörter und deren Nutzung für weitere Erwerbsprozesse nachgegangen. Unsere eigenen Untersuchungen zeigen, dass Säuglinge bereits in der zweiten Hälfte des ersten Lebensjahres unbetonte Funktionswörter im Input wahrnehmen, sie gegen Ende des ersten Lebensjahres als vertraute Einheiten erkennen und sie zur syntaktischen Kategorisierung anderer Wörter heranziehen. Damit verfügen Kinder in der ersten Hälfte des zweiten Lebensjahres bereits über maßgebliches syntaktisches Wissen.

In Kapitel 7 wird zusammenfassend versucht, auf Grundlage der Befunde ein genaues Bild der Sprachverarbeitung und des Erwerbs von Kenntnissen über die sprachliche Form während der ersten eineinhalb Lebensjahre zu zeichnen. Dabei wird vor allen Dingen deutlich, dass sich die frühe Entwicklung phonologischer und syntaktischer Aspekte der Zielsprache unter einer enormen Sensitivität des Säuglings für formale Aspekte seines sprachlichen Inputs vollzieht. Dabei verändert sich die Gewichtung der verschiedenen Informationsquellen und die Fähigkeit zur Integration verschiedener Informationsquellen während des ersten Lebensjahres beträchtlich, so dass der Säugling zu Beginn des zweiten Lebensjahres über Analysemechanismen verfügt, die ihm relativ zuverlässige strukturelle Analysen seines Inputs erlauben. Auf dieser Grundlage können Lernmechanismen strukturelle Regelmäßigkeiten der Zielsprache erkennen. In diesem Zusammenhang wird auch diskutiert, inwieweit der Erwerb struktureller Merkmale der Zielsprache diese Analysemechanismen selbst beeinflusst und deren Wirkungsweise für die spezifischen Eigenschaften der Zielsprache optimieren.

1 Grundlagen des Spracherwerbs

Unter normalen Bedingungen lernt jedes Kind im Laufe seiner ersten Lebensjahre mindestens eine Sprache zu verstehen und zu produzieren. Dieser Prozess verläuft im allgemeinen scheinbar mühelos und automatisch. Offensichtlich stellt ein ausreichender sprachlicher Input ohne eine explizite Unterweisung in der Grammatik der Zielsprache zumindest eine hinreichende externe Bedingung für einen erfolgreichen Spracherwerb dar. Zusätzlich muss das Kind jedoch auch über interne Voraussetzungen für den Spracherwerb verfügen, d. h. über bestimmte kognitive Fähigkeiten, die es ihm ermöglichen, die relevanten Merkmale seiner Umgebungssprache in relativ kurzer Zeit aus seinem sprachlichen Input herauszufiltern und das grammatische System der Zielsprache aufzubauen. Im Zentrum der jüngeren Spracherwerbsforschung steht nicht mehr die Frage, ob der Mensch überhaupt über diese besonderen internen Voraussetzungen für den Spracherwerb verfügt, sondern welcher Natur diese Voraussetzungen sind:

So the debate today on the field of language development is not about nature versus nurture, but about the „nature of nature,“ that is, whether language is something that we do with an inborn language device, or whether it is the product of (innate) abilities that are not specific to language.

Bates & Goodman, 1999, S. 33

Im Bereich des Spracherwerbs kommt dem Zusammenspiel angeborener Fähigkeiten bzw. angeborenen Wissens und der Erfahrung mit einem spezifischen, externen System, das seinerseits selbst wiederum dieses Wissen beeinflusst, eine besondere Rolle zu. In kaum einem anderen Bereich der kognitiven Entwicklung wird die Notwendigkeit einer Interaktion angeborener Fähigkeiten mit Umwelterfahrungen so deutlich wie im Bereich des Spracherwerbs. Schließlich differiert der Erwerbsgegenstand in Abhängigkeit von der jeweiligen Sprachgemeinschaft, in der das Kind aufwächst, deutlich. Da anzunehmen ist, dass die kognitiven Voraussetzungen, die den Spracherwerb ermöglichen für die gesamte Spezies identisch sind, müssen diese generell genug sein, um den Erwerb aller natürlichen Sprachen in ihrer gesamten Divergenz in gleicher Weise zu unterstützen. Diese Unterschiedlichkeit des Lerngegenstands scheint geradezu

nach Erklärungsansätzen zu verlangen, die die Erfahrung mit dem spezifischen sprachlichen System, in dem das Kind heranwächst, als Hauptmotor für die Sprachentwicklung begreifen. Eine Vielzahl von Beobachtungen spricht jedoch dafür, dass der Spracherwerb in weiten Teilen stärker von internen kognitiven Fähigkeiten und Reifungsprozessen vorangetrieben wird als von externen Bedingungen des sprachlichen Inputs.

So finden sich bei Kindern, die mit den verschiedensten Muttersprachen unter den verschiedensten Umweltbedingungen aufwachsen, identische Phasen im Spracherwerb: im allgemeinen durchlaufen alle Kinder präverbal die sogenannte Babbelphase, die ersten Äußerungen sind üblicherweise Einwortäußerungen, die ersten Mehrwortäußerungen zeigen übereinstimmend bestimmte Charakteristika wie fehlende Funktionsmorpheme und in vielen Sprachen die Verwendung infiniter Verbformen. Wenn das Alter, mit dem verschiedene Kinder diese Phasen erreichen und die Dauer der einzelnen Phasen auch interindividuell variiert, so findet sich jedoch wenig Variation in der Abfolge des Durchlaufens der einzelnen Phasen (s. z. B. Goodluck, 1991). Zudem existiert Evidenz für eine sogenannte kritische Phase für den Spracherwerb, die ungefähr bis zum Erreichen der Pubertät (Lenneberg, 1967) dauert. Nach dem Ende der kritischen Phase erscheint es nicht mehr möglich, anhand der gleichen Mechanismen, die dem Erstspracherwerb während der ersten Lebensjahre zugrundeliegen, eine volle muttersprachliche Kompetenz in einer Sprache zu erlangen oder wiederzuerlangen (als Überblick s. Epstein, Flynn & Martohardjono, 1996).

Darüber hinaus reflektieren Merkmale kindersprachlicher Äußerungen bzw. des sprachlichen Systems, das Kinder erwerben, nicht unbedingt direkt die Merkmale ihrer spezifischen sprachlichen Umgebung und ihrer spezifischen Erfahrungswelt. Zum einen finden sich nur geringe Korrelationen zwischen individuellen Charakteristika der Sprachverwendung zwischen Eltern und ihren Kindern (als Überblick s. Valian, 1999). Zum anderen zeigt sich, dass Kinder, die nur einen rudimentären sprachlichen Input erhalten, ein strukturell differenzierteres sprachliches System entwickeln können, als dies im Input erkennbar wäre. Ein Beispiel hierfür bietet die sogenannte Kreolisierung. Kreolsprachen entstehen dort, wo Kindern während des Erstspracherwerbs als Input lediglich sehr rudimentäre Pidginsprachen zur Verfügung stehen, ihr eigenes erworbenes sprachliches System, die Kreolsprache, dann jedoch strukturelle Merkmale

aufweist, die die zugrundeliegenden Pidginsprachen nicht enthalten (s. Bickerton, 1999). Ein analoges Phänomen findet sich auch in der Verwendung von Gebärden bei taubstummen Kindern, die nicht mit einer Gebärdensprache als Muttersprache aufwachsen, zu denen jedoch isolierte Gebärden verwendet wurden. Aus diesen isolierten Gebärden kreieren solche Kinder rudimentäre Regeln, zur Kombination der Gebärden, die auch bei Kindern unterschiedlicher Sprachgemeinschaften sehr ähnliche Charakteristika aufweisen (Goldin-Meadow & Mylander, 1998; Goldin-Meadow, Mylander & Butcher, 1995; Newport, 1988).

Die Antworten auf die Frage nach der Natur der angeborenen Voraussetzungen für den Spracherwerb lassen sich grob in zwei theoretische Ansätze differenzieren: den spezifischen Nativismus und den generellen Nativismus oder auch Konstruktivismus (s. z. B. O'Grady, 1999). Innerhalb des spezifischen Nativismus steht die Annahme im Vordergrund, dass das Kind über angeborenes spezifisch syntaktisches Wissen verfügt, eine sogenannte Universalgrammatik (Chomsky, 1965; 1981). Im Konstruktivismus herrscht dagegen die Auffassung vor, dass allgemeine domänenübergreifende Kategorien, Wahrnehmungsprinzipien und Lernmechanismen auch für den Erwerb des sprachlichen Systems ausreichen (Piaget, 1980; Bates & Goodman, 1999).

1.1 Der spezifische nativistische Ansatz

Wie schon gesagt gehen Vertreter des spezifischen Nativismus davon aus, dass der Mensch über angeborenes sprachspezifisches Wissen als eine der kognitiven Voraussetzungen für den Spracherwerb verfügt. Diese Universalgrammatik ist vor allen Dingen die Grundlage für den Erwerb syntaktischer Regularitäten der Zielsprache:

In the case of language, abstracting way from unknown mechanisms, we assume that the language faculty has an initial state, genetically determined, common to the species apart from gross pathology, and apparently unique to the human species.

Chomsky, 1999, S. 41

Die Universalgrammatik stellt eine Menge von angeborenen Prinzipien und Beschränkungen dar, die der Lerner in die Spracherwerbsaufgabe mit einbringt

(Chomsky, 1981). Diese Prinzipien und Beschränkungen bestimmen die möglichen sprachlichen Strukturen und somit auch die möglichen Annahmen des Lerners über die syntaktischen Einheiten und Regularitäten seiner Zielsprache. Innerhalb der sogenannten Kontinuitätshypothese (s. Lust, 1999) wird angenommen, dass dem Kind das universalgrammatische Wissen von Geburt an zur Verfügung steht, und dass sich dieses Wissen im Laufe des Spracherwerbs nicht verändert, d. h. alle syntaktischen Regularitäten, die direkt universalgrammatische Prinzipien reflektieren, sollte das Kind von Geburt an beherrschen. Scheinbare Veränderungen sprachlicher Fähigkeiten, die auf diesen Prinzipien beruhen, die sich beispielsweise im Bereich der Interpretation von Anaphern zeigen, werden als Ausdruck sich verändernder genereller Performanzbedingungen interpretiert (z. B. Crain, 1991). Dies bedeutet auch, dass der Erwerb einzelsprachspezifischer syntaktischer Regularitäten die Universalgrammatik selbst nicht berührt. Der Spracherwerb des Kindes besteht nach dieser Auffassung ausschließlich darin, die spezifischen Merkmale der jeweiligen Zielsprache, die universalgrammatisch aufgrund von zwischensprachlicher Variation nicht festgelegt sein können, zu identifizieren:

We know that this initial state can mature to a number of different steady states – the various attainable languages – as conditions of exposure vary. The process of maturation from the initial state to the steady state of mature knowledge is, to some extent, data driven; exposed to data of English, the mind/brain will incorporate knowledge of English, not Japanese. (.....) The environment determines how the options left undetermined by universal grammar are fixed, yielding different languages.

Chomsky, 1999, S. 41

Innerhalb des sogenannten Parameter-Fixierungs-Ansatzes (als Überblick s. Weissenborn, Goodluck & Roeper, 1992) wird ein Modell dafür geliefert, wie auf der Grundlage universalgrammatischen Wissens einzelsprachspezifisches Wissen erworben wird. Nach diesem Modell umfasst die Universalgrammatik bestimmte universelle Prinzipien als Parameter, die in verschiedenen Sprachen unterschiedliche Werte einnehmen können. So gilt beispielsweise die Tatsache, dass eine syntaktische Phrase eine Kopf- und eine Komplementierer-Position aufweist, als universalgrammatisches Merkmal, das der Phrasenstruktur aller Sprachen zugrundeliegt. Was jedoch zwischen Sprachen variiert, ist die Anordnung dieser Positionen innerhalb der Phrase: so gibt es

sogenannte kopffinale Sprachen, in denen der Kopf dem Komplementierer folgt und es gibt kopffinitiale Sprachen, in denen der Kopf dem Komplementierer vorangeht. Die Aufgabe des Kindes während des Spracherwerbs besteht darin herauszufinden, welche Werte die Parameter in der zu lernenden Sprache einnehmen, also die Parameter zu fixieren:

The initial state of the language faculty consists of a collection of subsystems, or *modules* as they are called, each of which is based on certain very general principles. Each of these principles admits of a certain very limited possibility of variation. We may think of the system as a complex network, associated with a switch box that contains a finite number of switches. The network is invariant, but each switch can be in one of two positions, on or off. (.....) For example, the possible phrase structures of a language are fixed by general principles and are invariant among languages, but there are some switches to be set. One has to do with order of elements. In English, for example, nouns, verbs, adjectives and prepositions precede their objects; in Japanese, the comparable elements follow their object. (.....) To acquire a language, the child's mind must determine how the switches are set, and simple data must suffice to determine the switch settings.

Chomsky, 1999, S. 49-50

Eine Voraussetzung für die Fixierung dieser Parameter besteht demnach darin, dass dem Kind spezifische Informationen aus dem sprachlichen Input der jeweiligen Zielsprache, sogenannte Trigger (z. B. Roeper & Weissenborn, 1990), verfügbar sind. Diese Trigger sollen sprachliche Merkmale sein, die dem Input auf direkte Weise zu entnehmen sind. Beispielsweise werden rhythmische Eigenschaften der Sprache als ein Trigger diskutiert, der es dem Kind erlaubt zu entscheiden, ob seine Zielsprache zu den kopffinitalen oder den kopffinalen Sprachen gehört. So sind phonologische Phrasen in kopffinitalen Sprachen auf dem letzten Element betont, weisen also prosodisch eine schwach-stark-Struktur auf, während phonologische Phrasen in kopffinalen Sprachen auf dem ersten Element betont sind, also eine stark-schwach-Struktur aufweisen (Nespor, Guasti & Christophe, 1996). Mit dem Erkennen des dominierenden Betonungsmusters in der phonologischen Phrase erwerben die Kinder nach diesem Ansatz gleichzeitig Wissen über die spezifischen Merkmale der Phrasenkonfiguration in ihrer Sprache und somit Grundlagen der Wortstellungsregeln für ihre Sprache.

Die Annahme einer Universalgrammatik als notwendige Voraussetzung für den Spracherwerb basiert in erster Linie darauf, dass gerade die Domäne der Syntax mit

ihren hochgradig spezifischen Beschränkungen und Prinzipien von einem generellen Lernmechanismus nicht zu erwerben sei. So existieren spezifische syntaktische Beschränkungen der Grammatik, die nicht aus offensichtlichen semantischen, pragmatischen oder diskursbezogenen Beschränkungen resultieren. Dies zeigt sich unter anderem daran, dass die Ungrammatikalität bzw. Grammatikalität einer Wortsequenz sich nicht unbedingt auf ihre Interpretierbarkeit auswirkt. Ein Beispiel hierfür sind die Beschränkungen bei Bewegungen oder die Bindungsprinzipien (aus Pinker, 1987):

- a. John saw Mary with her best friend's husband./Who did John see Mary with?
John saw Mary and her best friend's husband./ *Who did John see Mary and?

- b. John liked the pictures of Bill that Mary took./John liked Mary's pictures of Bill.
John liked the pictures of himself that Mary took./*John liked Mary's pictures of himself.

Die Beispiele zeigen, dass die Ungrammatikalität der jeweiligen Sätze kaum auf semantischen Beschränkungen beruhen kann, denn die ungrammatischen Sätze sind vor dem Hintergrund der grammatischen Strukturen durchaus interpretierbar. Die Beispiele zeigen aber darüber hinaus, dass ein unrestringierter Generalisierungsmechanismus zu falschen Ergebnissen kommt, denn die jeweils ersten Sätze der Satzpaare in (a) bzw. (b) sehen strukturell sehr ähnlich aus, da sie bis auf ein Wort identisch sind. Ein rein distributioneller Analysmechanismus sollte aufgrund dieser Verteilung zu dem Ergebnis kommen, dass die beiden Wörter, die in den ersten Sätzen der Paare jeweils differieren, zur gleichen syntaktischen Klasse gehören, so dass die gleichen Beschränkungen für diese Elemente gelten. Die ungrammatischen Sätze zeigen jedoch, dass sie nicht als strukturell identisch behandelt werden dürfen. Da dies erst über die Ungrammatikalität der entsprechenden Sätze erkannt werden kann, befinden wir uns in einem Zirkularitätsproblem.

Auch strukturell ambige Sätze machen das Problem der Erkennbarkeit von Strukturen an der Oberfläche eines Satzes deutlich. Strukturell ambige Sätze sind solche Sätze, die in der Abfolge der Wörter identisch sind, denen jedoch unterschiedliche Phrasenstrukturen zugrundeliegen, wie das vielzitierte Beispiel *Der Förster beobachtete*

den Mann mit dem Feldstecher. Aufgrund verschiedener struktureller Zuordnungsmöglichkeiten der Präpositionalphrase (entweder direkt vom Verb dominiert, oder in die Nominalphrase *den Mann* integriert) entstehen auch zwei unterschiedliche semantische Interpretationsmöglichkeiten für den Satz. Dies macht deutlich, dass sich selbst an einer grammatisch geordneten Kette von Wörtern, die zugrundeliegende syntaktische Struktur nicht direkt ablesen lässt.

Der Tatsache, dass die syntaktische Struktur einer Äußerung nicht direkt markiert ist, tragen auch Modelle der Sprachverarbeitung Rechnung, die annehmen, dass in die syntaktische Analyse eines Satzes Wissen über bestimmte strukturelle Prinzipien und Kategorien eingeht. So nimmt beispielsweise Frazier in ihrem Satzverarbeitungsmodell an (z. B. Frazier & Fodor, 1978; Frazier & Clifton, 1996) an, dass für die Berechnung der syntaktischen Struktur einer Äußerung die Identifizierung der syntaktischen Kategorie der in der Äußerung auftretenden Wörter ein erster grundlegender Schritt ist. Ohne die Identifikation der syntaktischen Kategorie der lexikalischen Elemente ist der Verarbeitungsmechanismus nicht in der Lage, beispielsweise die Phrasenstruktur einer Äußerung zu ermitteln. Dies muss bedeuten, dass ein Kind, das die Wörter einer Äußerung noch nicht syntaktisch klassifizieren kann, die syntaktische Struktur der Äußerung auch nicht erkennen kann. Damit befindet sich der Lerner in einem Dilemma (vgl. Pinker, 1984, 1987).

Neben diesen eher theoretischen Argumenten für die Annahme angeborenen sprachspezifischen Wissens für den Syntaxerwerb gibt es auch eine Reihe empirischer Argumente für diese Hypothese. Nach der Theorie des spezifischen Nativismus ist bei Störungen der kognitiven Entwicklung die Dissoziierbarkeit der Sprachentwicklung von anderen kognitiven Domänen zu erwarten, d. h. es sollte Sprachentwicklungsstörungen geben, ohne dass gleichzeitig eine andere, nicht-sprachliche kognitive Beeinträchtigung vorliegt, und es sollte Beeinträchtigungen der kognitiven Entwicklung geben, die nicht mit Sprachentwicklungsstörungen einhergehen. Für eine solche Dissoziierbarkeit der Entwicklung kognitiver Domänen gibt es eine Reihe empirischer Hinweise. So werden die sogenannten spezifischen Sprachentwicklungsstörungen als Beleg für eine isolierte Störbarkeit der sprachlichen Entwicklung herangezogen. Als spezifische Sprachentwicklungsstörung werden solche Störungen der Sprachentwicklung - primär der morphologischen und syntaktischen Entwicklung - bezeichnet, bei denen die

betroffenen Kinder keine weiteren offensichtlichen kognitiven, emotionalen oder behavioralen Beeinträchtigungen aufweisen (z. B. Clahsen, 1999). Eine isolierte Störung des Aufbaus gerade der morpho-syntaktischen sprachlichen Fähigkeiten würde sehr deutlich für domänenspezifische Voraussetzungen für den Grammatikerwerb sprechen.

Demgegenüber zeigen Kinder mit dem Williams-Beuren Syndrom, einer schweren, genetisch bedingten geistigen Beeinträchtigung, ein andersartiges Profil der kognitiven Entwicklung. Diese Patienten weisen relativ gute sprachliche Fähigkeiten auf, vor allen Dingen in den formalen Bereichen der Phonologie und der Morphosyntax, bei gleichzeitig massiven Beeinträchtigungen der Verarbeitung räumlicher und numerischer Information (z. B. Bellugi, Wang, & Jernigan, 1994; Clahsen & Almazan-Hamilton, 1998; Karmiloff-Smith, 1992). Damit deutet die Symptomatik der spezifischen Sprachentwicklungsstörungen und die des Williams-Beuren Syndroms auf eine doppelte Dissoziation zwischen sprachlichen Fähigkeiten und anderen kognitiven Domänen hin: die spezifische Sprachentwicklungsstörung zeigt Störungen der Sprachentwicklung bei unbeeinträchtigten non-verbalen Fähigkeiten, das Williams-Syndrom zeigt gute sprachliche Fähigkeiten bei massiven Beeinträchtigungen non-verbaler Domänen. In der kognitiven Neuropsychologie gilt ein solches Muster der doppelten Dissoziation als sehr starke Evidenz für die Unabhängigkeit kognitiver Domänen (z. B. Caramazza, 1984).

1.2 Der generelle Nativismus

Der generelle Nativismus beruht auf den Grundannahmen des Konstruktivismus, wie er von Jean Piaget (z. B. Piaget, 1980) vertreten wurde. Innerhalb des Konstruktivismus wird angenommen, dass es kein angeborenes domänenspezifisches Wissen gibt, sondern dass die kognitive Entwicklung allein auf der Grundlage von domänenübergreifenden Verarbeitungs- und Lernprozessen vorangetrieben wird. Im Bereich der Sprachentwicklung wird die Annahme, dass Spracherwerb ohne angeborenes sprachspezifisches Wissen erklärbar ist, in jüngerer Zeit unter anderem in den Arbeiten von Elizabeth Bates und Brian MacWhinney vertreten:

The universal properties of grammar are only indirectly innate, being based on interactions among innate categories and processes that are not specific to language. In other words we believe in the innateness of language, but we are skeptical about the degree of domain-specificity that is required to account for the structure and acquisition of natural languages.

(Bates & MacWhinney, 1989, S. 10)

Im Gegensatz zum universalgrammatischen Ansatz wird also angenommen, dass jegliches sprachliche Wissen auf der Grundlage angeborener nicht-sprachspezifischer Kategorien und Prozesse erworben wird. Der Gegensatz zum spezifisch nativistischen Ansatz wird besonders in der Auffassung über die Genese syntaktischer Kategorien und Relationen deutlich. Im spezifisch nativistischen Ansatz werden diese als Bestandteil der Universalgrammatik und somit als angeborenes sprachliches Wissen behandelt (Chomsky, 1981). Im generell nativistischen Ansatz wird demgegenüber angenommen, dass syntaktische Kategorien und Relationen ihren Ursprung in konzeptuellen sowie semantischen Kategorien und Relationen haben (Bowerman, 1973; Braine, 1992; Schlesinger, 1988), d. h. es wird angenommen, dass syntaktische Relationen wie Subjekt und Objekt und auch Wortklassen aus semantischen und letztlich konzeptuellen Rollen wie Agens, Patiens und Thema entwickelt werden:

It has been shown (...) that semantic categories, like agent, patient, location, etc., are sufficient for describing the very young child's knowledge of language and that there is no evidence of more abstract, syntactic categories in early speech.

Schlesinger, 1988, S. 122

If anything, it is these cognitively based categories that must be assumed to be innate.

ebd. S. 123

Neben diesen angeborenen generellen Kategorien werden ausschließlich domänenübergreifende Wahrnehmungs- und Verarbeitungsmechanismen, die unter anderem aus dem sprachlichen Input die für den Erwerb relevante Information herausfiltern, als ausreichend angesehen:

Language acquisition becomes a problem of pattern detection that may or may not require the application of innate linguistic knowledge. We suspect that more general principles of pattern detection and distributional learning are sufficient for the task.

(Bates & MacWhinney, 1989, S. 26)

Die Möglichkeit, sprachliche Regularitäten direkt aus dem sprachlichen Input zu induzieren, wurde in jüngerer Zeit vor allen Dingen innerhalb konnektionistischer Ansätze überprüft (s. Elman, Bates, Johnson, Karmiloff-Smith, Parisi & Plunkett, 1996). In diesem Zusammenhang kommt der Frage ein wesentlicher Stellenwert zu, welche direkt dem sprachlichen Input zu entnehmende Merkmale ein Lernalgorithmus nutzen kann, um strukturelle Merkmale zu erkennen und zu erwerben. Mittlerweile liegen zahlreiche Arbeiten vor, die zeigen, dass sprachliche Äußerungen eine ganze Reihe von formalen Hinweisen enthalten, die beispielsweise zur Kategorisierung sprachlicher Einheiten herangezogen werden können (z. B. Durieux & Gillis; 2001; MacWhinney, Leinbach, Taraban & McDonald, 1989). Allerdings wurde der Erfolg solcher rein inputbasierter Analysemechanismen und Lernmechanismen bislang ausschließlich in Computersimulationen gezeigt. Bei solchen Simulationen werden parallel relativ umfangreiche Korpora von natürlich-sprachlichen Äußerungen analysiert. Es ist fraglich, ob die parallele Verarbeitung größerer Textmengen die tatsächlichen Verarbeitungskapazitäten des Kindes reflektieren. In anderen Studien wurden sehr kleine, spezifische Korpora wie Wortlisten als Input verwendet. Auch diese konzentrierte Präsentation bestimmter Information entspricht sicherlich nicht der natürlichen Erwerbssituation des Kindes.

Eine kritische Frage, der sich der Ansatz des generellen Nativismus stellen muss, ist die nach einer Erklärung für die mögliche Dissozierbarkeit von sprachlicher und nicht-sprachlicher kognitiver Entwicklung, wie sie oben beschrieben wurde. Nach diesem Ansatz ist die Existenz einer spezifischen Sprachentwicklungsstörung auszuschließen. Bei jeder Sprachentwicklungsstörung sollten also auch Störungen in anderen kognitiven Domänen nachweisbar sein.

Tatsächlich gibt es für die Erklärung der Ursachen von Sprachentwicklungsstörungen auch eine Reihe von domänenübergreifenden Ansätzen. So schlägt beispielsweise Tallal (z.B. Tallal & Piercy, 1974; Tallal & Stark, 1981) ein grundlegendes Wahrnehmungsdefizit bei der Verarbeitung sehr kurzer oder sehr schnell

aufeinanderfolgender Reize als Ursache für Sprachentwicklungsstörungen vor. In verschiedenen empirischen Untersuchungen zeigte sich, dass Kinder mit Sprachentwicklungsstörungen sehr schnell aufeinanderfolgende Reize schlecht diskriminieren können und einen größeren zeitlichen Abstand zwischen den Reizen bzw. eine höhere Dauer der Reize selbst zur erfolgreichen Diskrimination benötigen als dies bei sprachlich unauffälligen Kindern der Fall ist. Dieses Muster zeigte sich sowohl für verbale als auch für nonverbale akustische Reize (Tallal & Piercy, 1973, 1974). Es stellt sich natürlich die Frage, warum dieses domänenübergreifende Wahrnehmungsdefizit in einer scheinbaren domänenspezifischen Sprachentwicklungsstörung mündet. Eine mögliche Antwort darauf könnte sein, dass gerade für die Verarbeitung des akustischen Signals gesprochener Sprache schnelle, effiziente Verarbeitungsmechanismen notwendig sind. Beispielsweise entspricht die Information, die zur Identifizierung einzelner Lautsegmente notwendig ist, zum Teil nur sehr kurzen zeitlichen Abschnitten im Signal. Der Teil des akustischen Signals, der die notwendige Information zur Unterscheidung zwischen stimmhaften und stimmlosen Verschlusslauten notwendig ist, umfasst Zeitabschnitte, die im allgemeinen nicht mehr als 100 ms umfassen. Das Wahrnehmungsdefizit würde aufgrund der Besonderheiten des Signals also sehr deutlich die Sprachwahrnehmung beeinträchtigen, nicht aber unbedingt die Wahrnehmung von Signalen, in denen sehr kurze Zeitabschnitte weniger relevant sind, so dass sich diese Störung möglicherweise ausschließlich negativ für den Spracherwerb auswirken könnte.

Im Zusammenhang mit dem Williams-Beuren-Syndrom wird verstärkt die Frage nach sprachlichen Beeinträchtigungen dieser Patientengruppe gestellt. Entgegen der oben dargestellten populären Auffassung, dass im Williams-Beuren-Syndrom sprachliche, zumindest morpho-syntaktische Fähigkeiten unbeeinträchtigt sind, mehren sich mittlerweile Hinweise darauf, dass diese Annahme nicht haltbar ist. So finden sich in einer Untersuchung von Karmiloff-Smith, Tyler, Voice, Sims, Udwin, Howlins und Davies (1998) Hinweise auf ein Defizit in der Verarbeitung von Subkategorisierungsinformation bei Patienten mit Williams-Syndrom. Darüber hinaus hatten die Patienten erhebliche Schwierigkeiten bei der korrekten Rollenzuweisung in semantisch reversiblen Sätzen, wie beispielsweise *The clown photographs the policeman*. In einer Satz-Bild-Zuordnungsaufgabe wählten sie überzufällig häufig das

Bild, auf dem eine vertauschte Rollenzuweisung abgebildet war. Über solche Vertauschungsfehler bei semantisch reversiblen Sätzen berichten auch Volterra, Capirci und Caselli (2001) in einer Untersuchung mit italienischen Patienten mit Williams-Beuren Syndrom. Dieses Fehlermuster gilt im allgemeinen als Hinweis auf spezifische Probleme bei der syntaktischen Analyse von Sätzen. Karmiloff-Smith et al. selbst interpretieren das Leistungsmuster der Williams-Beuren Patienten allerdings als Ausdruck eines Problems der Informationsintegration und nicht als Ausdruck einer speziell die Syntax betreffenden Störung. Diese Interpretation müsste jedoch mit weiteren Untersuchungen überprüft werden. Beim gegenwärtigen Stand der Forschung kann keine eindeutige Aussage darüber getroffen werden, ob Williams-Beuren-Patienten spezifische, die Syntax betreffende Störungen aufweisen oder nicht.

1.3 Spracherwerb und Sprachverarbeitung

Trotz dieser theoretischen Gegensätze gibt es einen gemeinsamen Nenner des spezifischen und des generellen Nativismus: beide Ansätze gehen davon aus, dass das Kind über effiziente Verarbeitungs- und Analysemechanismen für seinen sprachlichen Input verfügen muss, die es ihm erlauben, in relativ kurzer Zeit, die Information, die für den Erwerb der Zielsprache notwendig ist, aus seinem Input herauszufiltern. Die Untersuchung dieser Fähigkeiten wurde in der Spracherwerbsforschung lange Zeit weitgehend vernachlässigt. Dies mag auch methodische Ursachen haben, da die Untersuchung von Sprachverarbeitungsmechanismen auf die Verwendung experimenteller Methoden angewiesen ist, die zum Teil nur eingeschränkt mit Kindern verwendbar sind. Allerdings sind in den letzten Jahrzehnten eine Reihe experimenteller Techniken entwickelt worden, anhand derer bereits während der ersten beiden Lebensjahre detaillierte Untersuchungen zur Sprachwahrnehmung und -verarbeitung möglich sind (vgl. Kapitel 3).

1.3.1 Das System der Operating Principles

Slobin (1973) schlug ein System von Prinzipien (Operating Principles) vor, die die Art und Weise, wie Kinder ihren sprachlichen Input verarbeiten, speichern und analysieren, leiten. Diese Prinzipien können grob in zwei Kategorien eingeordnet werden, zum einen solche Prinzipien, die den sprachlichen Input in mentale Repräsentationen transformieren und zum anderen solche Prinzipien, die auf den gebildeten mentalen Repräsentationen operieren und diese weiter analysieren.

So wird beispielsweise die Erkennung und Speicherung linguistisch relevanter Einheiten wie Wörter und Phrasen durch bestimmte Prinzipien geleitet, die von Peters (1985) näher beschrieben werden. Hierbei geht es zunächst darum, wie aus dem sprachlichen Input Einheiten extrahiert werden. Peters schlägt vor, dass das Verarbeitungssystem hierzu alle Arten von Information nutzt, die bestimmte Teile des Inputs hervorheben bzw. abgrenzen, also beispielsweise prosodische Salienzmerkmale wie Betonung, Grenzsiknale wie Pausen, aber auch distributionelle Merkmale wie das wiederholte Auftreten von Sequenzen. Für die Einheiten, die nach diesen Prinzipien aus dem Input extrahiert wurden, werden mentale Repräsentationen gebildet. Über diese Repräsentationen werden zusätzliche Operationen ausgeführt, die einer weiteren Segmentierung dienen. So werden beispielsweise verschiedene Einheiten in ihrem phonologischen Aufbau verglichen: finden sich Einheiten mit phonologisch identischen Anteilen, werden sowohl die phonologisch identischen Anteile als auch die restlichen nicht-identischen als unabhängige, eigenständige Elemente gespeichert.

Zusätzliche Prinzipien tragen beispielsweise dafür Sorge, dass weitere im Input auftretende Token einer einmal gespeicherten Repräsentation zugeordnet werden und somit die Auftretenshäufigkeit einer Einheit ermittelt wird. Zudem wird Information über mögliche Kontexte der Einheiten und Kookkurrenzen zwischen verschiedenen Einheiten gesammelt, anhand derer die Repräsentationen syntaktisch klassifiziert und zu morphologischen Paradigmen zusammengefasst werden (Slobin, 1985).

Slobin (1985) charakterisiert diese Prinzipien als angeboren, demnach sind sie bereits vor jeglicher Erfahrung mit irgendeiner Art von sprachlichem Input verfügbar. Neben einem kleinen Anteil von generellen Problemlösungsstrategien, wie beispielsweise Generalisierungsmechanismen, sind die Prinzipien sehr spezifisch in Hinblick auf die Verarbeitung und Analyse sprachlicher Merkmale formuliert. Es bleibt

jedoch unklar, ob diese Prinzipien bereits in ihrem Ursprung ausschließlich auf die Domäne der Sprache ausgerichtet sind, oder ob sie aus generelleren Prinzipien der Wahrnehmung und Informationsverarbeitung abgeleitet sind und auf den besonderen Fall der Verarbeitung und Analyse sprachlicher Merkmale angepasst wurden.

1.3.2 Bootstrapping-Mechanismen

Ein Teil der von Slobin und Peters vorgeschlagenen Prinzipien finden sich in neueren Ansätzen zum sogenannten Bootstrapping wieder. Das Bootstrapping-Problem wie es am explizitesten von Pinker (1984, 1987) beschrieben wurde, besteht in der Frage, auf welche Weise das Kind beginnt, in seinem jeweiligen sprachlichen Input wesentliche Informationen aufzufinden. Selbst wenn die Existenz bestimmter syntaktischer Prinzipien und Kategorien dem Kind als angeborenes Wissen zur Verfügung steht, muss es in der Lage sein, dieses mit seinem sprachlichen Input in Verbindung zu bringen, d. h. es muss Wege finden, beispielsweise Exemplare verschiedener syntaktischer Kategorien in seinem einzelsprachlichen Input zu identifizieren. Innerhalb der Bootstrapping-Ansätze wird generell davon ausgegangen, dass das Kind für den Einstieg in den Erwerb einzelsprachspezifischen strukturellen Wissens Korrelationen zwischen unterschiedlichen sprachlichen Ebenen benutzt. Zentral ist dabei die Annahme, dass das Kind bereits Zugriff auf Informationen auf einer sprachlichen Ebene hat und diese Informationen nutzt, um Repräsentationen auf der korrelierenden Ebene aufzubauen (vgl. Weissenborn & Höhle, 2001). In Abhängigkeit von dem Informationstyp, den das Kind bereits ausnutzen kann, wird zwischen verschiedenen Arten von Bootstrapping unterschieden.

Prosodisches Bootstrapping

Innerhalb des Ansatzes des prosodischen Bootstrapping wird davon ausgegangen, dass das Kind akustische Merkmale im sprachlichen Signal als Hinweise auf die syntaktische Struktur der zugrundeliegenden Äußerung und damit auch auf das zugrundeliegende grammatische System nutzen kann (z. B. Gleitman & Wanner, 1982). In erster Linie sind in diesem Zusammenhang bestimmte Korrelationen zwischen prosodischen und syntaktischen Einheiten relevant, d. h. Korrelationen zwischen prosodischen Einheiten

wie beispielsweise Intonationsphrasen, phonologischen Phrasen und Füßen mit syntaktischen Einheiten wie Sätzen, Phrasen und Wörtern. Da die Grenzen prosodischer Einheiten häufig durch spezifische akustische Parameter wie beispielsweise bestimmte Dauer- oder Tonhöhenmerkmale markiert sind, wird angenommen, dass die prosodischen Einheiten sehr direkt im sprachlichen Signal identifiziert werden können. Über die Prosodie-Syntax-Korrelation wäre es dann möglich, einer Äußerung auch Informationen über ihre syntaktischen Einheiten und somit ihrer syntaktischen Struktur zu entnehmen (vgl. Cooper & Paccia-Cooper, 1980). Wie Morgan (1986) zeigt, ist das Lernbarkeitsproblem im Fall eines Inputs, der Informationen über Grenzen syntaktischer Phrasen enthält, weitaus geringer als für einen Input, der keinerlei Information über syntaktische Grenzen enthält.

Die mögliche Relevanz prosodischer Information wird darüber hinaus auch im Zusammenhang mit dem Erwerb von syntaktischen Kategorien für die Unterscheidung verschiedener Wortklassen diskutiert, sowohl für die Unterscheidung zwischen lexikalischen und funktionalen Elementen (Gleitman & Wanner, 1982) als auch für eine Differenzierung innerhalb der lexikalischen Elemente zwischen Nomen und Verben (Kelly, 1996). Zudem wurde vorgeschlagen, dass prosodische Information auch für den Erwerb von Wortstellungsregularitäten eine wichtige Informationsquelle darstellen könnten (z.B. Mazuka, 1996; Nespor, Guasti & Christophe, 1996).

Semantisches Bootstrapping

Pinker (1984, 1987) schlägt vor, dass Kinder das Problem der Identifikation syntaktischer Kategorien und Relationen initial durch Wissen über bestimmte Korrelationen zwischen semantischen und syntaktischen Kategorien lösen. Um diese Korrelation erfolgreich für die syntaktische Kategorisierung nutzen zu können, müssen sie lediglich dazu in der Lage sein, einzelnen Wörtern eine Bedeutung zuzuweisen und eine grobe semantische Repräsentation einer Äußerung zu erstellen. Die verfügbaren Korrelationen zwischen semantischen und syntaktischen Kategorien bestehen beispielsweise in dem Wissen, dass Objekt- und Personenbezeichnungen der syntaktischen Kategorie der Nomen entsprechen und dass Bezeichnungen für Handlungen oder Zustandsveränderungen Verben entsprechen. In ähnlicher Weise kann bei der Erkennung syntaktischer Relationen verfahren werden: so wird beispielsweise

die semantische Rolle des Agens auf die Subjektrelation abgebildet und die des Patiens auf die Objektrelation. In Verbindung mit angeborenem, konfigurationellem Wissen kann das Kind dieses erworbene syntaktisch-lexikalische Wissen dann im weiteren nutzen, um spezifisches Wissen über den Aufbau syntaktischer Phrasen in der Zielsprache aufzubauen.

Distributionelles Bootstrapping

Dieser Ansatz beinhaltet die Annahme, dass syntaktische Kategorien auf Grundlage distributioneller Information entdeckt werden können (z. B. Maratsos & Chalkley, 1980). Distributionelle Information umfasst hier sowohl Kookkurrenzen zwischen Wörtern verschiedener Klassen als auch Kookkurrenzen zwischen bestimmten Typen von Flexionsendungen und einer syntaktischen Klasse. So gibt es im sprachlichen Input gewisse positionale Relationen zwischen Wörtern bestimmter syntaktischer Kategorien: beispielsweise folgt einem Determinierer in vielen Fällen ein Nomen. In dieser Weise könnte also das Auftreten eines Determinierers für die Vorhersage der Wortklasse des folgenden Wortes verwendet werden. Tatsächlich konnten Mintz (1996) zeigen, dass die unmittelbaren lexikalischen Kontexte von Wörtern ein und derselben Wortklasse ähnlicher sind als die unmittelbaren lexikalischen Kontexte von Wörtern verschiedener Wortklassen. Die Beachtung eines lexikalischen Elements vor und hinter dem zu kategorisierenden Wort reicht bereits aus, um eine reliable Differenzierung zwischen Nomen und Verben zu erreichen.

Syntaktisches Bootstrapping

Im Rahmen des syntaktischen Bootstrappings wird angenommen, dass syntaktische Information zum Erwerb von Verbbedeutungen verwendet wird (Gleitman, 1990), speziell dass der Lerner den syntaktischen Rahmen, in dem ein Verb verwendet wird, dazu benutzt, um Information über die Argumentstruktur des Verbs zu gewinnen. Dabei spielen angeborene Korrelationen zwischen der syntaktischen Struktur einer Äußerung und der Argumentstruktur eines Verbs die ausschlaggebende Rolle. Tatsächlich zeigt experimentelle Evidenz, dass Kinder beispielsweise ein Verb, das ihnen in einem Satz mit Subjekt und Objekt geboten wird, als transitives Verb interpretieren und sie somit die Bedeutung des Verbs einschränken können (Naigles, 1990).

1.4 Zwei grundlegende Verarbeitungsschritte im frühen Spracherwerb: Segmentierung und syntaktische Kategorisierung

In der vorliegenden Arbeit steht die Auffassung im Mittelpunkt, dass Kinder bereits sehr früh über Verarbeitungsmechanismen verfügen, die sie zur Lösung des Segmentierungs- und des Kategorisierungsproblems heranziehen. Zudem wird die Kontinuität in den Verarbeitungsmechanismen bei Kindern und Erwachsenen angenommen, in dem Sinne, dass weder die Art der Information, die Kinder und Erwachsene aus ihrem sprachlichen Input nutzen, unterschiedlich ist, noch die Prozesse, mit denen diese Information verarbeitet wird (vgl. Mazuka & Lust, 1990). Die Grundannahme dabei ist, dass die gleichen Mechanismen und Informationstypen, die Kinder zum Aufbau sprachlichen Wissens nutzen, vom erwachsenen Sprecher für die Sprachverarbeitung genutzt werden (vgl. Christophe, Guasti, Nespore, Dupoux & van Ooyen, 1997). Ein wesentlicher Unterschied zwischen Kindern und Erwachsenen sollte lediglich sein, dass mit zunehmendem sprachlichen Wissen eine Zunahme der Relevanz wissensgetriebener *top-down* Prozesse für die Verarbeitung zu erwarten ist. So wird beim jungen Kind aufgrund des noch fehlenden einzelsprachlichen Wissens eine Dominanz rein inputgetriebener Verarbeitungsmechanismen erwartet, die im Laufe der Entwicklung zugunsten von wissensgetriebenen Prozessen abnimmt. Zudem wird angenommen, dass der Entwicklungsprozess von der anfänglich ausschließlichen inputgetriebenen Verarbeitung zum stärkeren Einfluss wissensgetriebener Prozesse kontinuierlich verläuft, d. h. das Kind nutzt sofort jegliches Wissen, das ihm zur Verfügung steht, sei es anhand seines Inputs erworbenes Wissen oder angeborenes Wissen für weitere Verarbeitungs- und Analyseprozesse des Inputs (vgl. auch Pinker, 1987).

Entsprechend der Forschungslage stehen in dieser Arbeit zum einen prosodische Merkmale des Inputs und deren Verarbeitung und zum zweiten nicht-prosodische distributionelle Information in Form von phonotaktischem und statistischem Wissen über Kookkurrenzen auf Phonem-, Silben- und Wortebene im Vordergrund. Im Sinne der Kontinuitätsannahme in Bezug auf die angenommenen Verarbeitungsmechanismen wird in jedem dieser Bereiche die Befundlage zur Verarbeitung dieser Informationstypen beim erwachsenen Sprecher/Hörer und die entsprechende

Befundlage beim Lerner gegenübergestellt, d. h. es werden jeweils drei Grundfragen gestellt:

- ?? Welche Informationen, die zu Segmentierung und Kategorisierung herangezogen werden können, enthält der sprachliche Input überhaupt?
- ?? Nutzen erwachsene Sprecher diese Merkmale für die Sprachverarbeitung?
- ?? Ab welchem Alter sind Kinder sensitiv für die entsprechende Information und nutzen sie diese Informationen zum Aufbau strukturellen Wissens?

1.4.1 Segmentierung des akustischen Signals

Das Segmentierungsproblem ergibt sich aus der Tatsache, dass das akustische Signal einer sprachlichen Äußerung kontinuierlich ist, d.h. es gibt keine eindeutigen akustischen Korrelate wie beispielsweise Pausen, die den Beginn oder das Ende eines Wortes markieren würden (z. B. Cutler, 1994; Shillcock, 1990). In vielen Ansätzen zum Spracherwerb wird trotzdem davon ausgegangen, dass der sprachliche Input, den das Kind verarbeitet, aus einer Kette von Wörtern besteht (Pinker, 1984; vgl. Mazuka, 1996). Der Repräsentation einer Äußerung als Kette von Wörtern müssen jedoch bereits Verarbeitungsschritte vorangehen, die das kontinuierliche Signal in wortentsprechende Segmente zerlegt, eine Aufgabe, deren Lösung keineswegs trivial ist. Studien zur Sprachverarbeitung machen deutlich, dass beim erwachsenen Sprecher neben Information aus dem Signal vorhandenes lexikalisches Wissen zur Lösung des Segmentierungsproblems beiträgt (Cole & Jakimik 1980; McQueen, Norris & Cutler, 1994). Es kann jedoch nicht davon ausgegangen werden, dass zu Beginn des Spracherwerbs lexikalisches Wissen in diesem Zusammenhang eine große Hilfestellung leisten kann: schließlich ist die Identifikation wortentsprechender Einheiten aus dem Input eine entscheidende Voraussetzung für den Aufbau lexikalischer Repräsentationen, die in ihrer Form den standardsprachlichen lexikalischen Repräsentationen entsprechen. Das Kind ist also stärker auf Informationen angewiesen, die es direkt seinem sprachlichen Input entnehmen kann. Offensichtlich hat der Lerner für das Segmentierungsproblem aber bereits innerhalb seines ersten Lebensjahres Lösungen gefunden, denn wenn Kinder beginnen, erste verstehbare Äußerungen zu produzieren,

was im allgemeinen am Ende des ersten Lebensjahres geschieht, so bestehen diese meistens zunächst aus einzelnen Wörtern, den sogenannten Einwortäußerungen.

Man könnte einwenden, dass für das Kind das Segmentierungsproblem gar nicht in dem Maße besteht, da sein Input über spezifische Merkmale verfügt, die das Kind zur Segmentierung nutzen kann. Tatsächlich zeigen genauere Studien kindgerichteter Äußerungen, dass das sogenannten *Motherese* eine Reihe von Besonderheiten aufweist, die das Segmentierungsproblem möglicherweise reduzieren. So sind kindgerichtete Äußerungen im allgemeinen kürzer als die an Erwachsene gerichteten, im Schnitt liegt die mittlere Äußerungslänge bei ca. drei Wörtern (Phillips, 1973; Bernstein Ratner & Rooney, 2001; van de Weijer, 1998). Vor allen Dingen ist natürlich in diesem Zusammenhang die Frage nach der Auftretenshäufigkeit von Einwortäußerungen im kindlichen Input von Relevanz. Hier schwanken die Zahlen in den Untersuchungen erheblich: Bernstein Ratner und Rooney (2001) fanden in ihrem kindgerichteten Korpus, der ca. 6000 an Kinder im Alter zwischen 13 und 20 Monaten gerichtete Äußerungen umfasste, 24% Einwortäußerungen. Bei van de Weijer (1998) zeigt sich, dass die Analyseergebnisse deutlich von dem zugrundegelegten Korpus abhängig sind: bei der Gesamtanalyse der an ein Kind zwischen dem 6. und dem 9. Lebensmonat gerichteten Äußerungen zeigte sich ein Anteil von annähernd 40% Einzelwortäußerungen. Wurden aus dem Korpus typische kindgerichtete Routinen wie beispielsweise *Kuckuck* oder *Hoppla* ausgeschlossen, sank die Rate der Einwortäußerungen auf 7%. Mit ähnlicher Vorgehensweise kommen Brent und Siskind (2001) bei einem ebenfalls bereinigten Korpus von Äußerungen, die an Kinder im Alter zwischen 9 und 15 Monaten gerichtet worden waren, zu analogen Ergebnissen. In ihren Daten zeigten sich 9% Einwortäußerungen.

Zwei weitere Merkmale kindgerichteter Sprache sind im Zusammenhang mit der Frage nach einer einfacheren Segmentierbarkeit relevant. Auf der einen Seite weisen kindgerichtete Äußerungen ein geringeres Type-Token-Verhältnis auf als an Erwachsene gerichtete Äusserungen, d. h. es finden sich mehr Wiederholungen der gleichen Wörter (Bernstein Ratner & Rooney, 2001; van de Weijer, 1998). Ein mehrfaches Auftreten der gleichen Phonemsequenz in einem geringen zeitlichen Abstand könnte ein Wiedererkennen dieser Sequenz und somit deren Extraktion aus dem Kontext begünstigen. Zudem zeigt kindgerichtete Sprache deutlichere prosodische

Merkmale als erwachsenengerichtete. So weist kindgerichtete Sprache in vielen Kulturen eine größere Variation der Tonhöhe, ausgeprägtere Intonationskonturen, sowie ein langsames Sprechtempo mit längeren Pausen auf (Fernald & Simon, 1984; Garnica, 1977; Grieser & Kuhl, 1988; Fernald, Taeschner, Dunn, Papousek, de Boysson-Bardies & Furui 1989; van de Weijer, 1998). Da gerade prosodische Merkmale als potentielle Hinweise auf Wort-, Phrasen- und Satzgrenzen gelten, ist kindgerichtete Sprache möglicherweise tatsächlich besser anhand reiner Inputmerkmale zu segmentieren als dies für erwachsenengerichtete Sprache der Fall ist.

Trotz dieser besonderen Merkmale kann nicht davon ausgegangen werden, dass die Eigenschaften der kindgerichteten Sprache den Säugling grundsätzlich vom Segmentierungsproblem befreien. Einerseits weist auch die kindgerichtete Sprache einen nicht unerheblichen Anteil von längeren Äußerungen auf. Zweitens sind die beschriebenen Merkmale nicht zwangsläufig universale Eigenschaften kindgerichteter Sprache (Ratner & Pye, 1984). Drittens zeigen bereits 8 Monate alte Säuglinge die Fähigkeit, einzelne Wörter aus kontinuierlicher Sprache, die nicht die auffälligen prosodischen Merkmale des *Motherese* aufweist, zu extrahieren (Saffran, Aslin & Newport, 1996). Dies deutet darauf hin, dass die Kinder über effiziente Verarbeitungsmechanismen verfügen, die es ihnen erlauben, auch im kontinuierlichen Sprachstrom Wortgrenzen aufzufinden.

1.4.2 Syntaktische Kategorisierung sprachlicher Einheiten

Innerhalb des spezifischen Nativismus stellt sich die Frage nach der Genese syntaktischer Kategorien nicht, da angenommen wird, dass die Universalgrammatik ein Inventar basaler lexikalischer und funktionaler Kategorien, wie beispielsweise Nomen, Verb etc. enthält (vgl. Chomsky, 1981, 1986).

Diesem Ansatz steht die konstruktivistische Auffassung gegenüber, nach der syntaktische Kategorien aus basalen konzeptuellen Kategorien entstehen. Wie bereits in Abschnitt 1.2 erwähnt, besteht der Kernpunkt dieses Ansatzes in der Annahme, dass dem Kind konzeptuelle Kategorien wie Agens, Patiens, Lokation von Geburt an zur Verfügung stehen. Anhand dieser Kategorien interpretiert das Kind seine Umwelt und lernt, wie diese Kategorien in der Zielsprache ausgedrückt werden. Demnach liegen

frühen kindersprachlichen Äußerungen keine syntaktischen, sondern ausschließlich semantische Relationen zugrunde: d. h. die sprachlichen Ausdrücke reflektieren direkt, ohne eine formalere, d. h. syntaktische Repräsentationsebene zu involvieren, diese konzeptuellen Relationen. Typische Zweiwortäußerungen wie beispielsweise *Mummy wash* sind daher nicht als Folge von Subjekt und Prädikat oder Nomen und Verb zu charakterisieren sondern als Agens-Aktion Abfolge.

Der konstruktivistische Ansatz der semantischen Quelle syntaktischer Einheiten bezieht sich im wesentlichen auf den Erwerb syntaktischer Relationen wie Subjekt, Prädikat und Objekt. Schlesinger (1982, 1988) beschreibt wie auf derselben Grundlage der Erwerb von Wortklassen zu erklären sei. Danach werden Wortklassen ebenfalls auf der Grundlage semantischer Relationen erworben, indem das Kind lernt, welche Wörter welche Argumente einer semantischen Relation besetzen können. Die Wörter, die die gleichen Argumentpositionen ausfüllen können, werden der gleichen Wortklasse zugeordnet:

For instance, the child learning English must discover that the word denoting the attribute precedes the one denoting the attributee. She has to extract this rule from input utterances like *big doggie, nice baby, dirty hands*, and so on. Now, as the child analyzes *big doggie* and notes how the semantic relation is realized, she will also register the fact that *big* is a word that can express the first argument of the attribute-attributee relation, i.e., the attribute, and *doggie* the second argument. In the course of hearing various realizations of this relation she thus comes to form two categories of words: words that can express the attribute argument (*big, nice, dirty*, etc.) and those that may serve for the expression of the attributees (*doggie, baby, hands*, etc.).

In the same manner the child will learn which words may serve to express other relational arguments: agents, patients, possessors, actions, and so on. Thus, learning the rules for realizing a semantic relation and learning which words are eligible for the expression of arguments of this relation are, on this view, parts of one and the same process.

Schlesinger, 1988, S. 169

Damit haben Wortklassen ihren Ursprung nicht in separaten konzeptuellen Kategorien wie Objekt, Person etc. sondern sind direkt an semantische Relationen gebunden. Ein Problem dieses Ansatzes besteht allerdings darin, wie Wörter, die nach syntaktischen Kriterien zur gleichen Wortklasse gehören, aber in unterschiedlichen Argumentpositionen auftreten, letztendlich in eine gemeinsame Klasse zusammengeführt werden. Dies ist ein notwendiger Schritt, da auch in diesem Ansatz

ein formal definiertes System syntaktischer Kategorien und Relationen das Erwerbsziel darstellt. Schlesinger (1988) schlägt hier folgende Lösung vor: Das Kind stellt im Laufe der Zeit fest, dass viele der erworbenen Wörter in verschiedenen Argumentpositionen der semantischen Relationen auftauchen können. Es führt dann alle Positionen, in denen identische Wörter auftauchen können in eine gemeinsame Kategorie über. Dies ist der Kern einer formal definierten Wortklasse. Wie der Autor selbst bemerkt, ist jedoch erstaunlich, warum Kinder manche Argumentpositionen in einer Kategorie bündeln und andere nicht: so können im Englischen häufig die gleichen Wortformen als Verb und als Nomen verwendet werden, trotzdem ist bei englisch lernenden Kindern keine Tendenz zu beobachten, Nomen und Verben als eine syntaktische Kategorie zu behandeln. Dies bedeutet jedoch auch, dass die Kinder andere Mittel zu Verfügung haben müssen, um syntaktische Kategorien zu bilden.

Ein weiterer Ansatz hält rein distributionelle Analysemechanismen für die Grundlage des Entstehens formaler sprachlicher Kategorien. Diese Annahme wurde vor allen Dingen von Maratsos in diversen Arbeiten (1982; 1988; Maratsos & Chalkley, 1980) vertreten:

The child must individually memorize that various individual predicates take *-ed* past tensing, various individual predicates take *-s* tensing, various individual predicates take preceding *do*-forms. (...) Over time the information must build up in long-term memory that there is a tendency for the same terms that take *-ed* past tensing to take *-s* tensing, to take preceding *do*-forms, and so on. This system of cross-implicational uses must form the basis for a process of category formation which operates when some groups of words share enough structural properties to justify forming a category. Once the category is formed, its shared distinctive properties then indicate membership in the category.

Maratsos, 1988, S. 34

Entscheidende Informationen zur Gruppierung von lexikalischen Elementen in eine Kategorie sind also gemeinsame Affixe von Wortformen und auch das Auftreten von Wortformen im selben Kontext. Da diese Eigenschaften auch genau jene sind, die gleiche Formklassen definieren, hält Maratsos sie für eindeutiger als beispielsweise eine anfängliche semantische Kategorisierung von lexikalischen Einheiten. Maratsos (1988) hält zwar einen konzeptuell-semantischen Einstieg in syntaktische Kategorien für theoretisch ebenso plausibel wie einen distributionsanalytischen, er führt jedoch eine

Reihe von Beobachtungen aus dem Spracherwerb an, die seiner Meinung nach gegen den konzeptuell-semantischen Ursprung formaler Kategorien sprechen. So sind die ersten Verben, die in kindersprachlichen Äußerungen auftauchen, nicht unbedingt Handlungsverben, sondern sie umfassen häufig auch Verben wie *see*, *want*, *need*, die nicht mit einer typischen Handlung assoziiert sind. Zweitens haben Kinder relativ wenig Probleme, semantisch nicht motivierte Systeme der Genuszuweisung zu erwerben (Levy, 1983). Für einen raschen Erwerb formaler Klassen spricht zudem die Tatsache, dass wortklassenverletzende Affigierungen kaum beobachtet werden. Alle diese Argumente sprechen nach Auffassung Maratsos dafür, dass formale Kategorien von Beginn an aufgrund formaler Übereinstimmungen zwischen mentalen Repräsentationen gebildet werden.

Analoge Mechanismen zur Gruppierung einmal aufgebaute lexikalischer Repräsentationen finden sich auch in Slobins (1973, 1985) Operating Principles wieder. Hier sind die Prinzipien relevant, die Kookkurrenzen verschiedener sprachlicher Elemente analysieren sowie nach phonologisch identischen Anteilen in phonologischen Repräsentationen suchen. Explizit für die syntaktische Klassifikation sind die folgenden Prinzipien relevant:

Store together as a class all words (phonological speech unit and meaning) that co-occur with a given functor. Store together as a class words that co-occur with the same groups of functors across utterances. Try to systematize word classes on semantic grounds, forming prototypes and looking for common features.

Slobin, 1985, S. 1252

Auch hier werden also zunächst Kookkurrenzmuster als Quelle für die Bildung formaler Kategorien verwendet, die jedoch dann anhand semantischer Kriterien systematisiert werden sollen. Auffällig ist in beiden Ansätzen zudem, dass im Zusammenhang der Etablierung von Wortklassen den funktionalen Elementen ein besonderer Stellenwert zugemessen wird: diese bilden die Kontexte, anhand derer die lexikalischen Kategorien etabliert werden.

Pinkers (1984, 1987) Hypothese des semantischen Bootstrapping basiert auf der Grundannahme, dass das Kind, ausgehend von semantischen Kategorien und Funktionen, syntaktische Kategorien und Funktionen in seinem sprachlichen Input

identifizieren kann. Im Gegensatz zu den oben dargestellten konstruktivistischen Ansätzen geht Pinker davon aus, dass die Existenz der syntaktischen Einheiten aber auch bestimmte Verknüpfungsregeln zwischen semantischen und syntaktischen Größen zum angeborenen universalgrammatischen Wissen des Kindes gehören:

Let me begin by pointing out what the semantic bootstrapping hypothesis does not claim. It does not claim that the young child lacks formal categories, or that his or her initial categories are organized around or defined by semantic criteria. Rather, the hypothesis presupposes that the child uses formal categories at all stages, and is intended to explain how the child knows which formal categories to posit in response to particular input sequences.

Pinker, 1984, S. 42

.. the claim of the Semantic Bootstrapping Hypothesis is that the child uses the presence of semantic entities such as „thing,“ „causal agent,“ „true in past,“ and „predicate-argument relation“ to infer that the input contains tokens of the corresponding syntactic substantive universals such as „noun,“ „subject,“ „auxiliary,“ „dominates,“ and so on.

Pinker, 1987, S. 407

Um diese Verknüpfungsregeln nutzen zu können muss das Kind also lediglich in der Lage sein, seinen sprachlichen Input semantisch zu interpretieren, d. h. vor allen Dingen semantische Relationen, wie Agens und Patiens zu identifizieren. Pinker ist der Auffassung, dass dem Kind dies mit Hilfe des nonverbalen Kontexts einer Äußerung und den Bedeutungen der einzelnen Wörter im Satz gelingen sollte. Sind durch diesen Verknüpfungsprozess erste Zuordnungen von sprachlichen Einheiten zu syntaktischen Kategorien geleistet worden, wird die semantisch basierte Kategorisierung durch eine strukturabhängige distributionelle Kategorisierung ergänzt.

An einigen Aspekten des semantischen Bootstrapping Ansatzes ist Kritik geübt worden. So bezweifelt Gleitman (1990, Gleitman & Gleitman, 2001) die Annahme, dass es dem Kind möglich ist, eine Äußerung allein anhand des Kontexts, in dem sie auftritt, zu interpretieren. Äußerungen kommentieren selten eine Handlung, die sich gerade vollzieht und beschreiben nicht unbedingt gerade die Situation, die sich innerhalb des Wahrnehmungsfeldes von Sprecher und Hörer abspielt. Außerdem gibt es für die meisten Situationen eine Vielzahl von Möglichkeiten, diese sprachlich zu beschreiben. Weiterhin wurde die Annahme universell gültiger Assoziationen zwischen bestimmten

semantischen und syntaktischen Einheiten bezweifelt. So werden beispielsweise in den sogenannten Ergativ-Sprachen eher thematische Rollen wie Thema oder Patiens als Subjekt ausgedrückt als das Agens (Marantz, 1984).

Der Begriff syntaktische Kategorisierung, wie er in der Spracherwerbsforschung verwendet wird, umfasst verschiedene Subprozesse. Das Kind muss über Wissen verfügen bzw. Wissen darüber erlangen, dass überhaupt verschiedene syntaktische Kategorien existieren. Es muss Merkmale ermitteln, die diesen syntaktischen Kategorien in der jeweiligen Sprache eigen sind. Es muss in der Lage sein, neu erworbene lexikalische Einheiten diesen Kategorien zuzuordnen. In der vorliegenden Arbeit geht es vorwiegend um den letztgenannten Prozess, nämlich um die Frage, welche formalen Hinweise in der Wortform selbst und welche distributionellen Hinweise im Prinzip nutzbar für die Zuordnung einer Wortform zu einer bereits etablierten syntaktischen Kategorie sind. Zudem steht die Frage nach der Evidenz im Vordergrund, dass Kinder bereits früh im Laufe des Spracherwerbs diese Informationen ausnutzen können.

2 Methoden zur Untersuchung früher sprachlicher und kognitiver Fähigkeiten

Lange Zeit konzentrierte sich die Spracherwerbsforschung im wesentlichen auf die Analyse spontaner Äußerungen von Kindern als primäre Datenquelle. Daraus ergab sich eine inhaltliche und altersmäßige Einschränkung auf Fragestellungen, die sich anhand der Merkmale sprachlicher Äußerungen untersuchen lassen. Dies führte dazu, dass der Spracherwerb während des ersten Lebensjahres weitgehend außer Acht gelassen wurde, da erste unstrittig als sprachlich betrachtete Äußerungen erst mit dem Beginn von Wortproduktionen, d. h. im allgemeinen gegen Ende des ersten Lebensjahres, zu beobachten sind. Erst in jüngerer Zeit richtet auch die produktionsorientierte Forschung ihr Augenmerk stärker auf das erste Lebensjahr. Diese Neuorientierung begründet sich auf der Erkenntnis, dass die vor allen Dingen in der zweiten Hälfte des ersten Lebensjahres auftretenden Babeläußerungen bereits spezifische Merkmale der Umgebungssprache aufweisen (De Boysson-Bardies, Hallé, Sagart & Durand, 1989; De Boysson-Bardies & Vihmann, 1991, Whalen, Levitt & Wang, 1991) und damit deutlich zeigen, dass sich bereits im ersten Lebensjahr wichtige Schritte für den Spracherwerb vollziehen.

Zudem zeigt eine Vielzahl neuerer Untersuchungen, dass die alleinige Analyse spontaner Produktionsdaten zu einer Unterschätzung des tatsächlichen sprachlichen Wissens, über das das Kind verfügt, führen kann. So verstehen Kinder weitaus komplexere Äußerungen, als sie selbst produzieren, und sie können Äußerungen, die in ihren formalen Merkmalen früher Kindersprache entsprechen - z. B. fehlende Funktionswörter im Telegrammstil - schlechter verarbeiten, als standardsprachliche grammatische Sätze (Gerken & McIntosh, 1993; Shipley, Smith & Gleitman, 1969). Um ein zuverlässigeres Bild über die altersmäßige Entwicklung sprachlichen Wissens zu erhalten, ist die Durchführung gezielter experimenteller Aufgaben mit den Kindern unerlässlich. Dabei kann es sich sowohl um Aufgaben handeln, die produktive Komponenten involvieren, beispielsweise Elizitierungsaufgaben, wie Nachsprechen, Beantwortung von Fragen etc., als auch um Aufgaben, die eine nonverbale Reaktion des Kindes erfordern, wie beispielsweise Satz-Bild-Zuordnungsaufgaben, Acting-Out-

Aufgaben etc. (für einen Überblick über gängige Paradigmen in der Spracherwerbsforschung siehe die Beiträge in McDaniel, McKee & Smith Cairns, 1996).

Die meisten dieser klassischen Untersuchungsparadigmen erfordern eine explizite Anweisung an das Kind, wie die Aufgabe durchzuführen ist, und setzen die Kooperationsbereitschaft des Kindes voraus. Zudem erfordert die Durchführung einer experimentellen Untersuchung die Schaffung einer formaleren Situation, als sie in einer normalen Spielsituation für das Kind alltäglich ist, da in den meisten Fällen beispielsweise von feedback oder zu häufigen Wiederholungen eines Stimulus abgesehen werden muss. Der Einsatz technischer Aufnahmegерäte wie Video- und Audiorekorder, die für eine zuverlässige Auswertung der Daten in den meisten Fällen unabdingbar ist, verstärkt den Charakter einer Testsituation. Es ist häufig schwierig, gerade für jüngere Kinder, eine Untersuchungssituation zu schaffen, die formal genug ist, um reliable Daten gewinnen zu können, aber nicht so formal, dass das Kind keine Kooperationsbereitschaft mehr zeigt. Aufgrund der verschiedenen Durchführungsprobleme sind die klassischen Aufgabenstellungen zur Untersuchung von Kindern unterhalb des dritten bis vierten Lebensjahres nur sehr eingeschränkt verwendbar.

Auf der anderen Seite zeigt sich übereinstimmend in vielen bisherigen Untersuchungen zum Spracherwerb, dass im morphosyntaktischen Bereich bis zum Erreichen des dritten bis vierten Lebensjahres wichtige Erwerbsschritte bereits abgeschlossen sind (s. z. B. die Arbeiten in Slobin, 1985). Aus der Tatsache, dass Kinder viele syntaktische Prinzipien bereits früh beherrschen und eine Entwicklung praktisch nicht nachweisbar ist, wurde von einigen Autoren der Schluss gezogen, dass das zugrundeliegende sprachspezifische Wissen angeboren sein muss (Crain, 1991). Dieser Schluss ist sicherlich nicht zwingend, wenn man bedenkt, dass zwei- bis dreijährige Kinder bereits einige Jahre Erwerbszeit hinter sich haben, - und wie man inzwischen vermutet, der Erwerb auch einzelsprachspezifischer Eigenschaften schon im Mutterleib beginnt (Jusczyk, 1997), da das Ungeborene bereits im letzten Drittel der Schwangerschaft Merkmale der außerhalb des Mutterleibs gesprochenen Sprache wahrnehmen kann (näheres s. Abschnitt 4.1.). Gerade um der Antwort auf die Frage, ob es angeborenes, spezifisch sprachlich strukturelles Wissen gibt, näher zu kommen, war

es notwendig, Methoden zu entwickeln, die eine möglichst frühe Untersuchung des sprachlichen Wissens sowie der Entwicklungs- und Lernprozesse in dieser Domäne ermöglichen.

Diese Lücke der ersten Lebensjahre – insbesondere der ersten beiden Lebensjahre – wird mittlerweile durch eine Reihe von erprobten experimentellen Techniken gefüllt, die mit Hilfe der Untersuchung von Sprachperzeption und Sprachverarbeitung, Erkenntnisse über den frühen Spracherwerb zulassen. Allen diesen Methoden ist gemein, dass sie ohne eine explizite Anweisung an den Probanden durchführbar sind, und entweder die spontane Reaktion auf einen bestimmten Reiz als abhängige Variable messen, oder nach einer Konditionierung der Versuchsperson die Stabilität der konditionierten Reaktion bei veränderten Stimulusbedingungen untersuchen. Im Rahmen der Spracherwerbsforschung werden hauptsächlich fünf Methoden verwendet, nämlich das sogenannte *High-Amplitude Sucking* Verfahren, das Verfahren der visuellen Fixierung, das *Headturn Preference* Paradigma, das *Conditioned Headturn* Paradigma sowie das *Preferential Looking* Paradigma. Diese Methoden werden im folgenden genauer diskutiert. Darüber hinaus gibt es weitere Paradigmen, die nur vereinzelt Verwendung gefunden haben (als Überblick siehe z.B. die Beiträge in Rovee-Collier, Lipsitt & Hayne, 1998).

2.1 Habituierungsparadigmen

2.1.1 Das High-Amplitude Sucking Paradigma

Innerhalb dieses Paradigmas wird als abhängige Variable die Stärke des Saugens des Kindes an einem speziell für die Messung dieses Parameters ausgestatteten Schnuller erhoben, während dem Kind ein akustischer Reiz präsentiert wird. Generell gilt, dass das Saugverhalten des Kindes mit seinem Interesse an einem visuellen oder akustischen Reiz korreliert, je interessanter dieser Stimulus für das Kind ist, umso höher ist die Saugrate. Aus diesem Grund ist das *High-Amplitude Sucking* Paradigma insbesondere zur Durchführung von Untersuchungen geeignet, in denen es um

Diskriminationsleistungen geht. Das Paradigma wurde zunächst für die Untersuchung der Verarbeitung visueller Stimuli entwickelt (Siqueland & Delucia, 1969), aber schon kurze Zeit später auch im Rahmen von Untersuchungen zur Sprachperzeption eingesetzt (Eimas, Siqueland, Jusczyk & Vigorito, 1971). Das Grundprinzip dieser Methode besteht darin, dass dem Kind nur dann ein Reiz präsentiert wird, wenn seine Saugstärke ein vorher definiertes Kriterium übersteigt. Offensichtlich sind Säuglinge sehr schnell in der Lage, den Zusammenhang zwischen ihrer Saugstärke und der Verstärkung durch die Präsentation eines visuellen oder akustischen Reizes zu erkennen (Siqueland & Delucia, 1969) und lernen so, die Stimuluspräsentation durch ihr Verhalten zu steuern. Bei einer Verstärkung durch immer denselben Reiz lässt die Saugstärke der Säuglinge nach, sie steigt jedoch wieder an, wenn zu einem neuen Verstärker übergegangen wird. Genau dieser Effekt wird für die Überprüfung von Diskriminationsleistungen ausgenutzt.

Eine Untersuchung mit dem *High-Amplitude Sucking* Verfahren besteht aus unterschiedlichen Phasen. In der ersten Phase erhält der Säugling nur den Schnuller ohne zusätzliche akustische Stimulierung. Dieses dient dazu, für jeden Säugling individuell sein Kriterium für einen *high-amplitude suck* zu bestimmen. Dazu lässt man den Säugling für eine bestimmte Zeitdauer am Sauger nuckeln und ermittelt anschließend die Werte des Drittels der stärksten Saugvorgänge. Der unterste Wert dieser 30% gilt dann als Schwellenwert für das Erreichen eines *high-amplitude sucks*. Während der eigentlichen Untersuchung wird dem Kind nur dann ein akustischer Reiz geboten, wenn seine Saugstärke über dem Schwellenwert liegt. Das Kind soll also während des Experiments einen Zusammenhang zwischen seinem Saugverhalten und der akustischen Stimulierung herstellen, in dem Sinne, dass starkes Saugen mit der Präsentation eines akustischen Reizes belohnt wird. In dieser Hinsicht enthält das *High-Amplitude Sucking* Paradigma auch eine Konditionierungskomponente.

Die eigentliche Untersuchung, in der dem Kind auch akustische Reize präsentiert werden, besteht aus zwei Phasen, der sogenannten Habituerungs- und der Dishabituerungsphase. Während der Habituerungsphase wird dem Kind bei jedem überschwelligen Saugen ein sprachlicher Reiz präsentiert, der während der Habituerungsphase entweder vollkommen oder in einem bestimmten Merkmal identisch bleibt (z.B. immer die Silbe /pa/). Typischerweise lässt die Häufigkeit des überschwelligen Saugens im Laufe der Habituerungsphase nach, da das Kind das

Interesse an dem sich stets wiederholenden Reiz verliert. Die Habituiierungsphase wird beendet, wenn die Häufigkeit des überschwelligen Saugens um ein vorher festgelegtes Kriterium nachgelassen hat. Im allgemeinen umfasst dieses Kriterium sowohl eine zeitliche Komponente als auch eine die Saugrate betreffende, beispielsweise ein kontinuierliches Absinken der Saugrate innerhalb von zwei aufeinanderfolgenden Minuten um 25% gegenüber der vorherigen Minute (z.B. Christophe, Dupoux, Bertoncini & Mehler, 1994). Ist dieses Habituiierungskriterium erreicht, wird in die Dishabituiierungsphase übergegangen. In dieser Phase wird einem Teil der Kinder, der Experimentalgruppe, ein gegenüber der Habituiierungsphase veränderter bzw. neuer Reiz präsentiert wird (z.B. die Silbe /ba/), während dem anderen Teil der Kinder, der Kontrollgruppe, der gleiche Reiz weiterpräsentiert wird. Ein signifikanter Anstieg der Saugrate bei der Experimentalgruppe bei konstanter Saugrate bei der Kontrollgruppe wird als Indikator dafür gewertet, dass die Kinder der Experimentalgruppe den Unterschied zwischen den beiden Reizen erkannt haben.

Die Methode wurde zunächst für einfache Laut- und Silbendiskriminationsaufgaben eingesetzt, in der während der Habituiierungs- und Dishabituiierungsphase lediglich jeweils ein einziger Stimulus dargeboten wurde (z. B. Eimas et al., 1971; Jusczyk & Thompson, 1978). Mittlerweile liegen jedoch auch Befunde aus Untersuchungen vor, in denen in beiden Phasen eine größere Bandbreite weitaus komplexerer Stimuli präsentiert wurden, die lediglich in einem kritischen Merkmal identisch waren. So konnte mit dieser Methode beispielsweise auch festgestellt werden, dass bereits Neugeborene zwischen Sprachen verschiedener Rhythmustypen unterscheiden können (Nazzi, Bertoncini & Mehler, 1998), dass wenige Tage alte Säuglinge Wörter mit verschiedenen Silbenzahlen unterscheiden können (Bijeljac-Babic, Bertoncini & Mehler, 1993) und dass zwei Monate alte Säuglinge Wortfolgeunterschiede in kurzen Sätzen registrieren (Mandel, Kemler Nelson & Jusczyk, 1996).

Das Paradigma eignet sich vor allen Dingen für Untersuchungen in den ersten Lebensmonaten. Es ist praktisch von Geburt an, aber nicht über das erste halbe Lebensjahr hinaus einsetzbar, da danach der Saugreflex deutlich nachlässt.

2.1.2 Visuelle Fixation

Die Methode der visuellen Fixation basiert auf der Beobachtung, dass Fixierungszeiten auf einen visuellen Reiz mit gleichzeitiger akustischer Stimulierung korrelieren, insbesondere werden die Fixierungszeiten länger, wenn der akustische Reiz verändert wird (Horowitz, 1975). Ebenso wie im *High-Amplitude Sucking* Paradigma ist die Stimuluspräsentation abhängig vom Verhalten des Kindes, hier von seiner Fixierung auf einen visuellen Reiz, im allgemeinen eine Abbildung eines geometrischen Musters oder eines Gesichts (s. z. B. Best, McRoberts & Sithole, 1988; Polka & Werker, 1994).

Während der Untersuchung befindet sich das Kind auf dem Schoß seiner Begleitperson in einer Kabine, die im allgemeinen mit einem Videomonitor zur Präsentation des visuellen Stimulus ausgestattet ist und einem Lautsprecher, der sich über oder unter dem Videomonitor befindet. In der Habituerungsphase der Untersuchung werden den Probanden in jedem Untersuchungsdurchgang die gleichen Reize dargeboten, beispielsweise eine konstante Silbe. Die Habituerungsphase dauert an, bis die Fixierungszeiten ein bestimmtes Kriterium unterschreiten. Dieses Kriterium variiert von Arbeit zu Arbeit. In der Untersuchung von Polka und Werker (1994) ist dieses Kriterium ein 50%iger Abfall der Fixierungszeiten bei zwei aufeinanderfolgenden Durchgängen gegenüber dem Mittelwert der beiden höchsten Fixierungszeiten der ersten drei Durchgänge. In dieser Arbeit wurde nach Erreichen dieses Kriteriums in den zwei folgenden Durchgängen ein anderer akustischer Stimulus geboten, nämlich eine andere Silbe. In den nächsten beiden Durchgängen wurde wiederum der Stimulus der Habituerungsphase geboten. Als Evidenz für eine erfolgreiche Diskrimination werden dabei signifikant längere Fixierungszeiten bei Präsentation des veränderten Stimulus gegenüber den beiden letzten Durchgängen der Habituerungsphase angesehen. Gleichzeitig dürfen sich jedoch die Fixierungszeiten der beiden letzten Durchgänge der Habituerungsphase nicht von denen der Testphase unterscheiden, in denen wiederum der akustische Reiz der Habituerungsphase präsentiert wird. Dieses Vorgehen macht den Einbezug einer Kontrollgruppe, denen nach Erreichen des Habituerungskriterium der gleiche akustische Reiz weitergeboten wird, überflüssig.

Wie andere Habituerungsparadigmen auch ist die visuelle Fixation im wesentlichen für die Untersuchung von Diskriminationsleistungen geeignet, bislang

standen dabei Untersuchungen zur Diskrimination von Phonemkontrasten im Vordergrund. Sie ist jedoch auch für die Untersuchung zur Diskrimination komplexerer Stimuli verwendbar, wie Untersuchungen von Pegg und Mitarbeitern (1992) zur Unterscheidung kindgerichteter und erwachsenengerichteter Sprache zeigen. Sie wurde erfolgreich bei Kindern ab dem vierten Lebensmonat eingesetzt. Über eine für dieses Paradigma spezifische obere Altersgrenze ist nichts bekannt.

2.2 Konditionierungsparadigmen

2.2.1 Das Conditioned Headturn Verfahren

Das klassische Verfahren

Die *Conditioned Headturn* Methode wurde alternativ zum *High-Amplitude Sucking* in erster Linie in Untersuchungen zur Lautdiskrimination verwendet. Alle Untersuchungen mit dem Paradigma bestehen aus einer Konditionierungs- und einer Testphase. In der Konditionierungsphase wird das Kind darauf trainiert, bei der Präsentation eines bestimmten sprachlichen Reizes den Kopf in Richtung eines seitlich angebrachten visuellen Verstärkers zu drehen. Bei korrekter Reaktion des Kindes, d. h. bei einer Kopfdrehung nach Präsentation des kritischen Reizes innerhalb eines bestimmten Zeitfensters, wird der visuelle Verstärker, häufig in Form eines beweglichen Spielzeugs für eine bestimmte Zeit in Gang gebracht.

Innerhalb des Experiments werden zwei verschiedene Stimulustypen präsentiert, der sogenannte Hintergrundstimulus und der eigentliche Teststimulus, auf den das Kind reagieren soll. Der Hintergrundstimulus wird in regelmäßigen Abständen wiederholt präsentiert, nur von Zeit zu Zeit wird in diesen Hintergrundstimulus eine Reihe von Teststimuli eingestreut. Viele der Untersuchungen zur kategoriellen Wahrnehmung und zur Lautdiskrimination bei Säuglingen wurden mit dem *Conditioned Headturn* Paradigma durchgeführt (z.B. Eilers, Wilson & Moore, 1977; Kuhl, 1979; Polka & Werker, 1994; Werker & Tees, 1984). In der Untersuchung selbst wird dabei als Hintergrundstimulus beispielsweise die Silbe /pa/ verwendet, während der

Konditionierung als kritischer Reiz eine bestimmte Variante der Silbe /ba/. Während der Testphase wird als kritischer Reiz eine andere Variante der Silbe /ba/ verwendet und getestet, ob auch diese vom Hintergrundstimulus diskriminiert wird, d. h. ob die konditionierte Reaktion der Kopfdrehung auch bei dem phonetisch veränderten Stimulus anhält.

Außerhalb von Fragestellungen zur Phonemdiskrimination liegen bislang meines Wissens keine erfolgreichen Untersuchungen mit diesem Paradigma vor. Es ist in seiner Anwendbarkeit relativ beschränkt, da eine Konditionierung auf einen einzelnen sprachlichen Stimulus erfolgen muss und in der Testphase auch lediglich die Reaktion auf diesen, bzw. einen leicht veränderten Stimulus gegenüber einem anderen Stimulus geprüft werden kann. D. h. innerhalb eines Experiments lässt sich stets nur eine Dimension im Reiz verändern und der Einfluss dieser einen Veränderung auf die Reaktion des Kindes überprüfen. Komplexere Fragestellungen lassen sich nur mit experimentellen Designs ohne Messwiederholung untersuchen, wobei mehrere Gruppen von Kindern untersucht werden müssen.

Das *Conditioned Headturn* Paradigma kann ab einem Alter eingesetzt werden, in dem die Säuglinge zu einer spontanen, selbständigen Kopfdrehung um mindestens 45° in der Lage sind. Im allgemeinen können sie dies im Alter von ungefähr 4 bis 5 Monaten. Prinzipiell gibt es keine obere Altersgrenze, es ist allerdings zu vermuten, dass bei älteren Kindern das Wechselspiel zwischen Lenkung der Aufmerksamkeit zum visuellen Verstärker und zum Untersuchungsleiter und den von ihm präsentierten Spielzeugen immer komplizierter wird.

Die Noise Detection Technik

Morgan (1994) entwickelte eine Variante des *Conditioned Headturn* Verfahrens, die einen Veteran psycholinguistischer Experimente, die sogenannten *click*-Experimente der 60iger Jahre (Fodor & Bever, 1965; Garrett, Bever & Fodor, 1965), für die Durchführung mit Säuglingen abwandelt. In diesem Verfahren werden die Kinder zunächst darauf trainiert, ihren Kopf bei der Präsentation eines Störgeräusches, beispielsweise eines Pieptones, in Richtung der Schallquelle zu drehen. In der eigentlichen Testphase der Untersuchung wird die Reaktion auf den Ton innerhalb verschiedener sprachlicher Kontexte untersucht. Als abhängige Variable lassen sich

dabei zwei verschiedene Parameter verwenden: zum einen die Häufigkeit von Treffern, d. h. korrekten Reaktionen des Kindes beim Auftreten des Tons und zum zweiten Latenzen, d. h. die Reaktionszeit, die zwischen der Präsentation des Störgeräuschs und der Initiierung der Kopfdrehung durch das Kind, liegt. Dieser zweite Parameter wurde in verschiedenen Untersuchungen von Morgan und Mitarbeitern (Morgan, 1996; Morgan & Saffran, 1995) verwendet, die der Frage nachgingen, unter welchen Bedingungen Kinder zweisilbige Sequenzen als kohärente Einheiten repräsentieren. Dazu wurden ihnen die Störgeräusche einmal innerhalb der Silbensequenzen präsentiert und einmal außerhalb. Gefundene Unterschiede in den Latenzen in Abhängigkeit von der Position des Störgeräuschs deuten beispielsweise darauf hin, dass Kinder rhythmisch strukturierte Sequenzen als kohärentere Einheiten wahrnehmen als rhythmisch nicht strukturierte.

Diese experimentelle Methode hat den Vorteil, dass der Stimulus, auf den die Kinder reagieren müssen, sich sehr deutlich von den sprachlichen Hintergrundstimuli unterscheidet. Damit scheint es auch möglich zu sein, den Einfluss von Merkmalen komplexerer sprachlicher Sequenzen auf die Detektion des Störgeräusches zu untersuchen. Gerade im Bereich des Syntaxerwerbs sind eine Vielzahl von Fragestellungen denkbar, die mit dieser Methode untersuchbar wären, da das Störgeräusch in Kontexte der verschiedensten Art eingebettet werden kann und man auf diese Art – im Sinne einer *dual task* – den Einfluss unterschiedlichster Kontexteigenschaften auf die Detektion des Störgeräuschs untersuchen kann.

Trotz dieser möglichen Anwendungsbreite ist diese Methode bislang verhältnismäßig wenig verwendet worden. Dies mag zum einen darin begründet sein, dass ihre Durchführung relativ aufwändig ist. So wurde in allen Untersuchungen, die Morgan selbst durchgeführt hat, die Konditionierung und die eigentliche Testphase auf zwei Tage verteilt, um die Dauer einer Sitzung in einem für junge Kinder angemessenen Bereich zu halten. In die eigentliche Testphase gehen darüber hinaus nur die Kinder ein, bei denen die Konditionierung erfolgreich verlaufen ist. Durch diese Beschränkungen ist die Anzahl der Ausfälle relativ hoch. Die für das klassische *Conditioned Headturn* Paradigma beschriebenen Schwierigkeiten der Durchführung der Experimente durch die Notwendigkeit der Ablenkung der Aufmerksamkeit vom visuellen Verstärker gelten natürlich auch hier. Auch die Altersgrenzen sind vergleichbar.

2.3 Präferenzparadigmen

2.3.1 Das Headturn Preference Paradigma

In diesem Paradigma wird als abhängige Verhaltensvariable die Dauer einer Kopfdrehung (Orientierungszeit) gemessen, die das Kind zur Fixierung eines seitlich angebrachten Lautsprechers macht, während von diesem Lautsprecher ein akustischer Reiz präsentiert wird. In die Forschung zum frühen Spracherwerb wurde das Paradigma in Untersuchungen von Fernald und Mitarbeitern eingeführt (Fernald, 1985; Fernald & Kuhl, 1987). In diesen ersten Experimenten war die Vorgehensweise allerdings etwas anders als heute üblich. Es wurde nicht die Orientierungszeit als abhängige Variable gemessen, sondern die Anzahl der Kopfdrehungen in die eine oder die andere Richtung. Dabei gab es eine feste Zuordnung von Stimulustyp und Präsentationsseite, die die Kinder in einer kurzen Konditionierungsphase zunächst lernen sollten. Durch ihre initiale Kopfdrehung zu Beginn eines Untersuchungsdurchgangs konnten die Kinder auswählen, welche Art von Stimulus sie hören wollten, da der Stimulustyp für die Seite präsentiert wurde, in die das Kind seinen Kopf drehte. In einer Untersuchung von Hirsh-Pasek und Mitarbeitern (Hirsh-Pasek, Kemler Nelson, Jusczyk, Cassidy, Druss & Kennedy, 1987) zeigte sich jedoch, dass die Menge der Kopfdrehungen ein weniger reliables Maß für die Präferenzen für einen Stimulustyp darstellen als die Orientierungszeit, so dass in gängigen Untersuchungen nur noch die Orientierungszeit in die Datenanalyse eingeht.

Untersuchungen mit diesem Paradigma werden in einer Untersuchungskabine durchgeführt, in der drei verschiedene Lampen angebracht sind: eine an der Rückwand der Kabine, in Blickrichtung des Kindes und zwei an den seitlichen Wänden, die in etwa an der Sitzposition des Kindes angebracht sind, so dass eine Kopfdrehung um ca. 90 Grad nötig ist, um eine dieser Lampen vollkommen zu fixieren. Hinter den beiden Seitenlampen sind außerhalb der Kabine, d. h. für das Kind nicht sichtbar, Lautsprecher für die akustische Stimulierung befestigt. Der Ablauf eines Untersuchungsdurchgangs sieht im allgemeinen so aus: Zunächst blinkt die an der Rückwand der Kabine angebrachte Lampe, um die Aufmerksamkeit des Kindes zu zentrieren. Wenn das Kind diese Lampe fixiert, wird sie gelöscht. Gleichzeitig beginnt eine der beiden seitlichen

Lampen zu blinken und zwar auf der Seite, auf der der nächste akustische Stimulus präsentiert wird. Wenn das Kind seinen Kopf in Richtung der jetzt blinkenden Lampe dreht und diese fixiert, wird der akustische Reiz gestartet. Von diesem Zeitpunkt an wird die Orientierungszeit gemessen. Die akustische Stimulierung wird so lange fortgesetzt, wie das Kind in Richtung Schallquelle schaut. Dreht das Kind seinen Kopf länger als 2 Sekunden weg, um in eine andere Richtung zu schauen, wird der Untersuchungsdurchgang abgebrochen und zum nächsten übergegangen. Auch in diesem Paradigma ist eine Konditionierungskomponente enthalten. Das Kind soll den Zusammenhang zwischen der Fixierung der seitlichen Lampe und der Dauer der akustischen Stimulierung erkennen, um zu erreichen, dass das Kind bei Stimuli, die es interessanter findet, eine längere Kopfdrehung aufrecht erhält. Dafür werden ihm zunächst einige Untersuchungsdurchgänge geboten, die nicht in die eigentliche Auswertung eingehen.

In die Auswertung geht die Dauer der Orientierungszeiten für die verschiedenen Typen von Stimuli, die im Experiment verwendet worden sind, ein. Dabei ist zu beachten, dass ein Experiment aufgrund der nachlassenden Aufmerksamkeit der Säuglinge während der Durchführung nicht mehr als ungefähr 16 Untersuchungsdurchgänge umfassen kann, so dass die Anzahl der Untersuchungsbedingungen im allgemeinen auf zwei eingeschränkt ist.

Untersuchungen mit diesem Paradigma können aus einer reinen Testphase bestehen, in der Wissen überprüft wird, das das Kind bereits mit in die Untersuchung einbringt. So wurde beispielsweise anhand dieses Verfahrens überprüft, ab wann Kinder grammatische von ungrammatischen Sätzen unterscheiden können (Höhle, Weissenborn, Schmitz & Ischebeck, 2001; Santelmann & Juszyk, 1998), anhand welcher Merkmale sie ihre Muttersprache von einer Fremdsprache unterscheiden (Juszyk, Friederici, Wessels, Svenkerud & Juszyk, 1993) und ab wann sie phonotaktische Regularitäten ihrer Muttersprache erkennen (Friederici & Wessels, 1993).

Neben der ursprünglichen Variante des *Headturn Preference* existiert eine zweite von Juszyk und Aslin (1995) entwickelte Methode, in der zwei experimentelle Phasen - eine sogenannte Familiarisierungsphase und eine Testphase - durchgeführt werden. In der Familiarisierungsphase und der Testphase werden den Kindern

unterschiedliche Typen von sprachlichen Stimuli bzw. ein sprachlicher Stimulus unter verschiedenen Bedingungen präsentiert. Diese Variante zielt weniger darauf ab, bereits vorhandenes sprachliches Wissen zu überprüfen, sondern sie überprüft Lern- und Verarbeitungsmechanismen, über die die Kinder im untersuchten Altersbereich verfügen. Von Jusczyk und Aslin wurde sie ursprünglich verwendet, um Segmentierungsprozesse zu untersuchen. Dazu wurden den Kindern während der Familiarisierung isolierte Wörter präsentiert und in der Testphase Textpassagen, die die zuvor präsentierten Wörter entweder enthielten oder nicht. Aus den unterschiedlichen Orientierungszeiten auf diese verschiedenen Texttypen lässt sich schließen, dass die Kinder die zuvor isoliert gehörten Wörter in der kontinuierlichen Sprache der Textpassage wiedererkannten.

Das *Headturn Preference* Verfahren ist mittlerweile das Verfahren, das in Untersuchungen zum frühen Spracherwerb am häufigsten verwendet wird. Mit der Möglichkeit, vorhandenes Wissen und auch Lern- und Verarbeitungsprozesse zu untersuchen, ergibt sich eine Fülle von Anwendungsbereichen dieses Verfahrens, die nicht auf bestimmte linguistische Domänen beschränkt sind, sondern sowohl in der Forschung zum Phonologie-, dem Lexikon-, und auch dem Syntaxerwerb relevant sind. Ein zweiter großer Vorteil dieser Methode liegt darin, dass das Altersspektrum der zu untersuchenden Kinder sehr breit ist. Die untere Altersgrenze ist im Prinzip von zwei Faktoren bestimmt: die Kinder müssen in der Lage zu einer selbständigen Kopfdrehung sein. Mit leichten Veränderungen im Aufbau der Untersuchungskabine lassen sich bereits Kinder im Alter von 4 Monaten mit dieser Methode untersuchen. Eine obere Altersgrenze gibt es prinzipiell nicht, bislang liegen Erfahrungen mit Kindern im Alter bis zu ca. 24 Monaten vor.

2.3.2 Preferential Looking

In dieser Methode wird die Fixationszeit des Kindes auf zwei simultan nebeneinander präsentierte visuelle Stimuli – je nach Fragestellung der Untersuchung Videosequenzen oder statische Objektabbildungen - bei gleichzeitiger Präsentation eines auditiven Reizes gemessen (Spelke, 1979). Das *Preferential Looking* Paradigma ist analog zu den üblichen Satz- oder Wort-Bild-Zuordnungsaufgaben aufgebaut, indem lediglich einer

der beiden visuellen Reize inhaltlich zum sprachlichen Reiz passt, während der andere einen Ablenker darstellt. Genau wie mit Satz- und Wort-Bild-Zuordnungsaufgaben lassen sich anhand dieser Methode Wort- und Satzverständnis überprüfen. Der Unterschied gegenüber diesen klassischen Methoden besteht darin, dass von den Kindern keine explizite Reaktion in Form der Auswahl eines Bildes verlangt wird, sondern lediglich das spontane Verhalten des Kindes auf die Präsentation des visuellen Reizes in Verbindung mit dem sprachlichen Reiz beobachtet wird. Als abhängige Variable wird dabei die Dauer der Fixation der visuellen Stimuli gemessen.

Diese Methode ist dazu geeignet, Verständnisse für den sprachlichen Stimulus zu untersuchen, denn generell zeigen Kinder längere Fixationszeiten für den visuellen Stimulus der zum sprachlichen passt, als für den Ablenker (s. Hollich, Hirsh-Pasek & Golinkoff, 1998). Nutzt man diese Methode beispielsweise für eine Überprüfung des Verständnisses von Passivsätzen - z. B. der Junge wird von dem Mädchen geschubst -, würde ein visueller Stimulus die im Satz genannte Situation darstellen, während der Ablenker die umgekehrte Relation, nämlich einen das Mädchen schubsenden Jungen, darstellen würde. Im Fall des korrekten Verständnisses des Passivsatzes wären längere Fixationszeiten für die passende visuelle Darstellung zu erwarten. Zur Kontrolle, ob Unterschiede in der Fixationszeit tatsächlich mit dem gleichzeitig präsentierten auditiven Stimulus zu tun haben, oder auf einer generellen Präferenz eines der beiden visuellen Reize beruhen, werden die Videosequenzen meist zunächst ohne gleichzeitigen akustischen Reiz geboten und die Fixationszeiten gemessen. Das Verfahren wurde bislang erfolgreich in Untersuchungen zur Entwicklung des Wortverständnisses als auch des Satzverständnisses angewendet (z. B. Golinkoff et al., 1987; Hirsh-Pasek & Golinkoff, 1996; Naigles, 1990).

Neben einer reinen Überprüfung des Verstehens von Wörtern lässt sich das Verfahren auch zur Untersuchung von lexikalischen Verarbeitungsprozessen nutzen. In dieser Variante ist nicht die Dauer der Fixation der visuellen Stimuli die abhängige Variable, sondern die Latenz vom Beginn des kritischen akustischen Reizes, beispielsweise der Anfang des kritischen Wortes, bis zur Initiierung einer Kopfdrehung zum korrekten visuellen Stimulus bzw. weg vom visuellen Stimulus. Damit ist ein *online*-Maß gegeben, das beispielsweise Auskunft über die Schnelligkeit von Worterkennungprozessen gibt (z. B. Swingle, Pinto & Fernald, 1998, 1999; Swingle

& Aslin, 2000). Die Methode kann in einem relativ breiten Altersspektrum eingesetzt werden. Erfahrungen liegen bislang im wesentlichen mit Kindern im Alter von ca. 15 bis 30 Monaten vor (s. z. B. Hirsh-Pasek & Golinkoff, 1996).

2.4 Familiaritäts- oder Neuigkeitseffekt: Die Richtung von Präferenzen

Studien, die die Dauer der spontanen Aufmerksamkeit, die einem Reiz zugewendet wird, erfassen, können neben reinen Diskriminationsleistungen auch Präferenzen für einen der präsentierten Stimulustypen ermitteln. In den meisten der frühen Untersuchungen mit dem *Headturn Preference* Paradigma zeigen Säuglinge eine Präferenz für die vertrauteren Stimuli, einen sogenannten Familiaritätseffekt. So fand man beispielsweise, dass Säuglinge Stimuli ihrer Muttersprache gegenüber denen einer fremden Sprachen bevorzugen (Jusczyk et al., 1993), dass sie Stimuli, die in der Muttersprache häufiger sind gegenüber solchen, die seltener sind, präferieren (z. B. Jusczyk, Cutler & Redanz, 1993; Jusczyk, Luce & Charles-Luce, 1994), dass grammatische Sätze gegenüber ungrammatischen Sätzen präferiert werden (Santelmann & Jusczyk, 1998) oder im selben Experiment schon mehrfach präsentierte Stimuli gegenüber noch nicht präsentierten Stimuli (z. B. Jusczyk & Aslin, 1995).

In jüngerer Zeit mehren sich jedoch Untersuchungen, in denen sich eine Präferenz für einen neuen, weniger vertrauten Stimulus, ein sogenannter *Novelty*-Effekt zeigt. Dies gilt sowohl für Familiarisierungsexperimente, in denen eine Präferenz für den neuen gegenüber einem vertrauteren Reiz gefunden wurde (Echols, Crowhurst & Childers, 1997; Johnson & Jusczyk, 2001; Saffran, 2001; Saffran, et al., 1996), als auch für Untersuchungen mit grammatischen und ungrammatischen Stimuli (Höhle et al., 2001), sowie für Stimuli, in die künstliche Pausen an nicht adäquaten Positionen eingefügt wurden, gegenüber Stimuli mit Pausen an syntaktisch adäquaten Positionen (Fernald & McRoberts, 1996).

Generell lässt sich also beobachten, dass bei der Präsentation sprachlicher Stimuli im *Headturn Preference* Paradigma Präferenzen sowohl für die vertrauten, als auch für die neuen Stimuli auftreten können. Für ein gegebenes Experiment scheint eine Voraussage, in welche Richtung ein Effekt zu erwarten ist, äußerst schwierig. Dies mag

zunächst als großer Nachteil dieser Methode erscheinen, für die Feststellung von Diskriminationsleistungen ist jedoch die Richtung einer Präferenz nicht von Bedeutung. Trotzdem ist diese Unsicherheit unbefriedigend und es bleibt ein wesentliches methodisches Desiderat, die Faktoren, die sich auf die Richtung von Präferenzen auswirken, näher zu bestimmen. In der Literatur finden sich Überlegungen auf mögliche relevante Einflussgrößen fast ausschließlich in Bezug auf die visuelle und die nicht-sprachliche akustische Domäne (s. Burnham & Dodd, 1998; Hunter & Ames, 1988; Wagner & Sakovitz, 1986). Einige der dort diskutierten Faktoren lassen sich jedoch auf Untersuchungen mit sprachlichen Stimuli übertragen.

Ein sehr explizites Modell, das eine Vorhersage des Auftretens von Familiaritäts- oder Neuigkeitseffekten ermöglichen soll, wurde von Hunter und Ames (1988) vorgeschlagen. Das Modell bezieht sich auf Untersuchungen, die aus einer Familiarisierungs- und einer Testphase bestehen, lässt sich jedoch auch auf Untersuchungen ohne Familiarisierungsphase übertragen. Das Modell nimmt für die Testphase des Experiments vier Phasen der Präferenz an: in der initialen Phase sollten sich noch keine Unterschiede zwischen dem familiarisierten und dem neuen Stimulus zeigen, da das Kind beide Stimulustypen zunächst explorieren muss, um Unterschiede festzustellen. Dieser Phase folgt eine Phase der Präferenz für den familiarisierten Stimulus, die dann in eine Präferenz für den neuen Stimulus umschlägt. Während dieses Umschlagens der Präferenz ist wiederum für einen gewissen Zeitraum kein Unterschied in der Aufmerksamkeit für die beiden Stimulustypen nachzuweisen.

Ein wesentlicher Punkt des Modells ist die Annahme, dass dieser Phasenverlauf nicht statisch ist, sondern dass die Schnelligkeit mit der die verschiedenen Phasen durchlaufen werden, in Abhängigkeit von unterschiedlichen Faktoren variiert. Eine der Größen, die in dieser Hinsicht relevant ist, ist das Alter der Probanden. Generell zeigen sich Tendenzen, dass bei zunehmendem Alter die Phasen schneller durchlaufen werden und sich somit in einer Untersuchung mit älteren Kindern mit höherer Wahrscheinlichkeit ein Neuigkeitseffekt zeigen sollte als bei der gleichen Untersuchung mit jüngeren Kindern.

Eine weitere Einflussgröße stellt nach Hunter und Ames (1988) die Aufgabenkomplexität dar, wobei eine zunehmende Aufgabenkomplexität den Ablauf der einzelnen Phasen verzögert und somit bei hoher Aufgabenkomplexität eher eine

Präferenz für die vertrauten Reize und damit ein Familiaritätseffekt zu erwarten ist. Die Aufgabenkomplexität ist ihrerseits wiederum ein komplexes Konstrukt, das von verschiedenen Merkmalen der im Experiment verwendeten Stimuli und der Untersuchungsdurchführung bestimmt wird. Hierbei ist zunächst die Komplexität der Stimuli entscheidend. Ein komplexer, schwerer zu verarbeitender Stimulus erhöht die Aufgabenkomplexität, so dass bei Untersuchungen mit komplexeren Stimuli mit höherer Wahrscheinlichkeit Familiaritätseffekte zu erwarten sind als bei Aufgaben mit Stimuli, die weniger Anforderungen an die Verarbeitungsmechanismen stellen. Auch Experimente, in denen die in der Familiarisierungs- und der Testphase verwendeten Stimuli sehr unterschiedlich sind, haben eine relativ hohe Aufgabenkomplexität, da während des Experiments zwei völlig verschiedene Typen von Stimuli zu verarbeiten sind. Auch in diesem Fall erhöht sich die Wahrscheinlichkeit eines Familiaritätseffekts gegenüber Aufgaben, in denen sich die Stimuli der Familiarisierungs- und der Testphase sehr ähnlich sind. Drittens soll der Grad der Ähnlichkeit der unterschiedlichen Stimuli während der Testphase modellgemäß einen Einfluss auf die Geschwindigkeit ausüben, mit der die verschiedenen durchlaufen werden. Sind die Stimuli der Testphase schwer zu diskriminieren, dauert möglicherweise die erste Phase der Exploration der Stimuli relativ lange, so dass eine Verzögerung der Phasen auftritt, die zu einer Erhöhung der Wahrscheinlichkeit für einen Familiaritätseffekt führt. Darüber hinaus kann die Dauer der Familiarisierung und somit die Häufigkeit der Präsentation des zu familiarisierenden Reizes wirksam sein. Bei einer häufigen Präsentation des Familiarisierungsstimulus wird durch die wiederholte Verarbeitung das Verarbeitungssystem für diesen Reiz habituiert, d. h. die Verarbeitungsanforderungen für diesen Reiz lassen nach. Dies führt in der Testphase zu einer insgesamt relativ geringen Aufgabenkomplexität, so dass hier eher mit einem Neuigkeitseffekt zu rechnen ist.

Während im Bereich der visuellen und der non-verbale auditiven Wahrnehmung gezielte Untersuchungen den Einfluss dieser genannten Faktoren auf das Auftreten von Familiaritäts- und Neuigkeitseffekten überprüften und dabei empirische Evidenz für das Modell erbrachten (z.B. Colombo & Bundy, 1983; Hunter, Ames & Koopman, 1983; Rose, Gottfried, Melloy-Carminar & Bridger, 1982), wurde innerhalb der Forschung zur Sprachperzeption das Problem der gegensätzlichen Richtung von

Effekten bislang kaum systematisch untersucht. Divergenzen in den Effekten einiger sehr ähnlicher Untersuchungen lassen jedoch darauf schließen, dass die diskutierten Einflussgrößen auch bei sprachlichem Stimulusmaterial wirksam sind. Hier sollen exemplarisch Unterschiede in zwei Studien, die bei sehr ähnlicher Vorgehensweise und Fragestellung zu gegenteiligen Effekten gelangten, diskutiert werden. Es handelt sich dabei um die Arbeiten von Jusczyk, Houston und Newsome (1999, Experiment 2) und Saffran et al. (1996), die anhand des *Headturn Preference* Paradigmas mit Familiarisierungsphase untersuchten, welche Information Säuglinge aus ihrem sprachlichen Input heranziehen, um Wortgrenzen zu erkennen. In beiden Untersuchungen wurde den Säuglingen während der Familiarisierungsphase kontinuierliche Sprache dargeboten. Die Stimuli der Testphase waren Teile aus dieser Familiarisierungssequenz und als neue Stimuli Silbenabfolgen, die in der Familiarisierungssequenz gar nicht (Jusczyk et al.) oder seltener (Saffran et al.) vorgekommen waren. Während sich in der Jusczyk et al.-Studie ein Familiarisierungseffekt zeigte, fanden Saffran und Mitarbeiter einen Neuigkeitseffekt. Welche Unterschiede zwischen den beiden Studien könnten für diese gegensätzlichen Ergebnisse relevant sein?

Erstens verwendeten Jusczyk und Mitarbeiter für die Familiarisierungsphase in ihrem Experiment zwei verschiedene Textpassagen, die aus sechs Sätzen bestanden, in denen das zu familiarisierende Wort jeweils einmal auftauchte. Bei einer 45 Sekunden dauernden Familiarisierung für jede Textpassage ist davon auszugehen, dass die Kinder jedes Zielwort insgesamt ca. 12 mal hörten. In der Testphase wurden die beiden familiarisierten Wörter und zwei neue nicht-familiarisierte Wörter geboten. Demgegenüber verwendeten Saffran und Mitarbeiter vier verschiedene dreisilbige Sequenzen sinnloser Silben, die in der Familiarisierungsphase als kontinuierlicher Sprachstrom in verschiedenen Reihenfolgen geboten wurden. Die Familiarisierungsphase dauerte insgesamt 2 Minuten, wobei jede dreisilbige Sequenz ca. 45mal vorkam. In der Testphase wurden zwei der dreisilbigen Sequenzen mit Sequenzen kontrastiert, die zwar in dem kontinuierlichen Silbenstrom auch adjazent vorgekommen waren, aber Silben verschiedener Dreisilbler beinhalteten. In dieser Beschreibung finden sich verschiedene der oben diskutierten Einflussgrößen wieder. Modellgemäß weist das Experiment, in dem sich der Neuigkeitseffekt zeigte, die

längere Familiarisierungsphase mit der häufigeren Präsentation des kritischen Stimulus auf. Zudem sind die in der Familiarisierungsphase der Jusczyk et al.-Studie verwendeten Textpassagen allein in formalen Aspekten variabler und – da sie natürliche Sprache mit prosodischen Merkmalen darstellen - informationsreicher als die von Saffran et al. verwendeten Silbensequenzen. Dadurch sind die Stimuli der Studie von Jusczyk und Mitarbeitern komplexer, was wiederum zu einer höheren Aufgabenkomplexität und somit eher zu dem tatsächlich beobachteten Familiarisierungseffekt führen sollte.

Der Vergleich dieser beiden Untersuchungen zeigt also, dass ein Einfluss der für die visuelle Modalität relevanten Einflussgrößen auch in der sprachlichen Modalität wahrscheinlich ist. Für die Formulierung von Erwartungen in Bezug auf die Richtung eines Effekts in einem Experiment zum Spracherwerb kommt jedoch erschwerend hinzu, dass sich die Aufgaben- und Stimuluskomplexität im Laufe des Spracherwerbs durch das zunehmende sprachliche Wissen des Kindes verändert, also auch hier eine Interaktion mit dem Alter des Kindes zu erwarten ist. Dabei ist nicht von vornherein gegeben, dass mit fortschreitendem Spracherwerb grundsätzlich die Komplexität einer Aufgabe nachlässt. Man kann sich sogar den gegenteiligen Effekt vorstellen. Während beispielsweise junge Kinder eine Textpassage vielleicht ausschließlich auf phonologischer Ebene analysieren und verarbeiten, findet bei älteren Kindern mit bereits besseren Sprachverständnisseleistungen wahrscheinlich gleichzeitig die Verarbeitung phonologischer, lexikalischer, syntaktischer und semantischer Information statt. Es ist denkbar, dass dies bei älteren Kindern sogar zu einer höheren Aufgabenkomplexität führt, als wenn das gleiche Stimulusmaterial mit jüngeren Kindern verwendet wird.

Über das genaue Zusammenspiel der verschiedenen Faktoren und eventuelle Besonderheiten bei Experimenten mit sprachlichen Stimuli ist bislang so gut wie nichts bekannt, so dass weiterhin mit der Unsicherheit über den Ausgang eines Experiments gelebt werden muss. Dieses ist kein Problem, wenn sich in einem Experiment signifikante Unterschiede zwischen den vertrauten und den weniger vertrauten Stimuli zeigen, da sich dieses unabhängig von der Richtung als Indikator für die Diskrimination der verwendeten Stimulustypen interpretieren lässt. Anders sieht es nach diesen Modellüberlegungen jedoch dann aus, wenn sich in einem Experiment kein

signifikanter Unterscheid zwischen den Stimulustypen zeigt. Neben den statistischen Problemen, die sich generell mit der Interpretation von Nulleffekten ergeben, taucht in diesen Paradigmen ein weiteres Problem auf: es ist möglich, dass sich ein anfänglicher Familiaritätseffekt und ein folgender Neuigkeitseffekt durch die übliche Mittelung über die gesamten Untersuchungsdurchgänge aufheben und somit zu einem scheinbaren Nulleffekt führen. Hier wird deutlich, dass aus der Tatsache, dass in einer Untersuchung keine signifikanten Unterschiede zwischen den verschiedenen Stimulustypen gefunden werden, auf keinen Fall geschlossen werden kann, dass diese von Kindern des untersuchten Alters nicht diskriminiert werden können.

3 **Prosodische Information und Sprachverarbeitung**

Bevor auf einen etwaigen Nutzen prosodischer Information – worunter hier Rhythmus, Betonung und Intonation verstanden wird (vgl. z. B. Goldsmith, 1990) - für den Spracherwerb eingegangen wird, soll in den folgenden Abschnitten diskutiert werden, welche Rolle diese Art von Information in der Sprachverarbeitung von Erwachsenen spielt. Wie schon im einleitenden Kapitel erwähnt, soll hier auch die Hypothese überprüft werden, dass sich bei Erwachsenen und Kindern weitgehende Gemeinsamkeiten in der Art der Information, die in der Sprachverarbeitung ausgenutzt wird, finden. Über die Verarbeitung inputbezogener Merkmale liegen für Erwachsene mittlerweile eine Vielzahl von Erkenntnissen vor, so dass es sinnvoll erscheint, diese als Ausgangspunkt auch für die Betrachtung des Spracherwerbs heranzuziehen.

Die Rolle prosodischer Information in der Sprachverarbeitung ist bei Erwachsenen insbesondere in zwei Domänen untersucht worden: zum einen im Zusammenhang mit Fragen der Worterkennung, speziell mit der Frage, auf welche Weise der Hörer das Signal in Einheiten segmentiert, die als Grundlage für den lexikalischen Zugriff dienen können. Den zweiten zentralen Forschungsgegenstand stellt die Rolle der prosodischen Struktur für die Berechnung der syntaktischen Struktur einer Äußerung dar, hier speziell bei der Erkennung von syntaktischen Phrasen und der Bestimmung der Wortklasse von lexikalischen Einheiten.

Generell wird in beiden Forschungsdomänen die Annahme gemacht, dass prosodische Information sehr direkt dem Signal zu entnehmen ist, so dass hier eine primär datengesteuerte Verarbeitung des Signals wichtige Informationen für die Verarbeitung auf höheren sprachlichen Ebenen – wie der lexikalischen und der syntaktischen Ebene – zur Verfügung stellen kann. Auch wenn diese Annahme – wie später noch zu diskutieren sein wird – nicht unproblematisch ist, muss innerhalb dieses Rahmens geklärt werden, welche Eigenschaften des sprachlichen Signals für diese datengesteuerte Verarbeitung relevant sein könnte.

3.1 Lexikalische Verarbeitung

Sehr gut untersucht ist die Rolle prosodischer, speziell rhythmischer Information im Rahmen der Frage, wie Hörer Wortgrenzen im akustischen Signal identifizieren können. Obwohl das akustische Signal gesprochener Sprache ein lautliches Kontinuum darstellt, in dem Wortgrenzen nicht in einheitlicher Weise markiert sind (vgl. z. B. Cutler, 1994), gehen alle gängigen Modelle der lexikalischen Verarbeitung (als Überblick s. z. B. Lively, Pisoni & Goldinger, 1994) davon aus, dass ein Lexikoneintrag einer isolierten Wortform entspricht. Um einen lexikalischen Zugriff zu ermöglichen, muss also irgendwann im Lauf der Verarbeitung des Signals ein Punkt gefunden werden, von dem aus die lexikalischen Suchprozesse gestartet werden können. Vor allen Dingen innerhalb des Kohortenmodells zur Worterkennung (Marslen-Wilson, 1987, 1990) wird in diesem Zusammenhang den initialen Segmenten eine entscheidende Rolle zugewiesen. Wenn auch eine vollständige prälexikalische Segmentierung des akustischen Signals in wortentsprechende Einheiten für eine schnelle und effiziente Worterkennung keine notwendige Voraussetzung ist, so wäre ein rasches Erkennen potentieller Wortanfänge jedoch für die Effizienz des Systems von Vorteil.

3.1.1 Rhythmische Segmentierung als Teil der lexikalischen Verarbeitung

Vor allen Dingen haben sich die Arbeitsgruppen von Anne Cutler und Jacques Mehler in vielen Studien der Frage nach prälexikalischen Segmentierungsmechanismen gewidmet. Nach ihrer metrischen Segmentierungshypothese sind die Grenzen der rhythmischen Grundeinheiten einer Sprache als prälexikalische Segmentierungspunkte ausschlaggebend. Ausgehend von der klassischen typologischen Differenzierung zwischen akzent-, silben- und morenzählenden Sprachen (Abercrombie, 1967) vermuten sie sprachspezifische Segmentierungsmechanismen, die in silbenzählenden Sprachen Silbengrenzen als wesentliche Segmentierungspunkte heranziehen, in akzentzählenden Sprachen dagegen nur die Grenzen vor betonten Silben, d. h. Fußgrenzen. Im dritten Sprachtypus, dem Japanischen, in dem die More die rhythmische Grundeinheit darstellt, sollen dagegen Morengrenzen die relevanten Segmentierungspunkte sein (Otake, Hatano, Cutler & Mehler, 1993).

Tatsächlich zeigte sich in einer Reihe von Untersuchungen, dass die nach diesem Ansatz vorhergesagten Grenzen als Segmentierungspunkte relevant sind: bei Sprechern der silbenzählenden Sprachen Französisch und Spanisch fanden sich deutliche Vorteile bei der Erkennung ganzer Silben gegenüber Phonemsequenzen, die keiner Silbe entsprachen (Bradley, Sánchez-Casas & García-Albea, 1993; Mehler, Dommergues, Frauenfelder & Segui, 1981; Pallier, Sebastián-Gallés, Felguera, Christophe & Mehler, 1993; Sebastián-Gallés, Dupoux, Segui & Mehler, 1992); Sprecher der akzentzählenden Sprachen Englisch, Deutsch und Niederländisch zeigten dagegen keinen entsprechenden generellen Vorteil bei der Erkennung ganzer Silben gegenüber nicht-silbischen Phonemsequenzen, für Sprecher dieser Sprachen scheint jedoch die Grenze vor einer betonten Silbe eine besondere Rolle für Segmentierungsprozesse zu spielen (Bradley et al., 1993; Cutler, Mehler, Norris & Segui, 1986; Finney, Protopapas & Eimas, 1996; Höhle & Schriefers, 1995; Zwitserlood, Schriefers, Lahiri & van Donselaar, 1993). Sprecher des morenzählenden Japanischen erwiesen sich sensitiver für Morengrenzen als für Silbengrenzen (Otake et al., 1993).

Interessanterweise sind diese unterschiedlichen Segmentierungspunkte offenbar keine direkte Reflektion dessen, dass in den verschiedenen Untersuchungssprachen Silbengrenzen mit unterschiedlicher Deutlichkeit akustisch markiert sind. So könnte man beispielsweise vermuten, dass im Französischen Silbengrenzen im allgemeinen akustisch besser markiert sind als im Englischen oder im Deutschen, in denen in einer Vielzahl von Wörtern unklare Silbengrenzen mit ambisyllabischen Konsonanten angenommen werden (vgl. z. B. Wiese, 1996). Untersuchungen, in denen den Probanden fremdsprachliches Material präsentiert wurde, zeigen jedoch, dass sie ihre muttersprachliche Segmentierungsstrategie auch anwenden, wenn ihnen Stimulusmaterial einer Sprache dargeboten wird, die einem anderen Rhythmustyp zugerechnet wird. So nutzen französische Sprecher im Gegensatz zu englischen Muttersprachlern eine Silbensegmentationsstrategie, auch wenn ihnen englische Sprache präsentiert wird, obwohl englische Muttersprachler keine Tendenz zu einer Silbensegmentation ihrer Muttersprache zeigen. Englische Muttersprachler zeigen zudem im Gegensatz zu französischen Muttersprachlern keine Silbensegmentationsstrategie bei der Verarbeitung französischer Stimuli (Cutler et al., 1986; Cutler, Mehler, Norris & Segui, 1992). Französische Muttersprachler

segmentieren auch japanische Stimuli in Silben und nicht – wie Sprecher des Japanischen – in Moren (Otake et al., 1993).

Diese Befunde deuten an, dass die Heranziehung verschiedener Segmentationspunkte auf sprachspezifischen Verarbeitungsmechanismen der Sprecher der verschiedenen Sprachen beruhen. Es stellt sich nun natürlich die Frage, auf welche Weise sich diese spezifischen Segmentierungsstrategien herausbilden. Cutler und Mitarbeiter (1992) vermuten, dass die gemeinsame Wurzel der verschiedenen Segmentierungsstrategien in ihrer Abhängigkeit von der rhythmischen Struktur der Sprache zu suchen ist. Sie gehen davon aus, dass eine rhythmusbasierte Segmentierungsstrategie zum angeborenen Repertoire des Sprachlerner gehört. Im Sinne einer Parameterfixierung wird diese Strategie im Laufe des Spracherwerbs an die rhythmischen Gegebenheiten der Zielsprache angepasst, d. h. beim französisch-lernenden Kind bildet sich beispielsweise die silbenbasierte Strategie heraus, beim deutsch-lernenden Kind die akzentbasierte.

In welcher Weise liefern diese Strategien nun Segmentierungen des sprachlichen Signals, die für die lexikalische Verarbeitung relevant sind? Die günstigste Ausgangslage für die weitere lexikalische Verarbeitung des Signals würde von einer Segmentierungsstrategie geschaffen, die zumindest in der Lage ist, die Wortanfänge im akustischen Signal mit hoher Wahrscheinlichkeit korrekt zu identifizieren. Diese Frage ist vor allen Dingen für das Englische untersucht worden. Nach der akzentbasierten Segmentierungsstrategie wäre der Beginn jeder betonten Silbe für das System ein potentieller Segmentierungspunkt und für die weitere lexikalische Verarbeitung damit auch ein potentieller Wortanfang. Cutler und Carter (1987) formulieren die metrische Segmentierungsstrategie als Algorithmus, der neben der Segmentierung des sprachlichen Inputs auch eine Kategorisierung in Inhalts- und Funktionswörter vornimmt, die auf deren metrischen Eigenschaften beruht :

- 1.1. *Lexical Structure.* The main lexicon contains only open-class (lexical) Lexical Structure
- 1.2. *Segmentation.* An initial segmentation process scans the input and places markers at the onset of each strong syllable.

1.3. *Lexical Access.*

- 1.3.1. If the initial string of the current input is not preceded by a marker, it is submitted to the closed-class list. If there is a marker, the string is submitted to the main lexicon.
- 1.3.2. The lookup process in both the main lexicon and the closed-class list returns the longest candidate consistent with the input, except that the occurrence of a marker indicating the beginning of a strong syllable will terminate the current lookup process, returning the longest word found so far (if any), and initiate a new lookup process in the main lexicon. When a match is found, the process returns to the input (i.e. starts again at 1.3.1.)
- 1.3.3. Failure of either lookup process leads to the same string being submitted to the other store.
- 1.3.4. Failure in both stores leads to backtracking, i.e. cancellation of a previous decision. This may involve accepting a shorter candidate word, if there is one; or undoing a word assignment for the preceding syllable and attaching it to the current input string; or over-riding a following marker and initiating a lookup process for a string with a medial strong syllable. (We have no evidence on which to base a proposal for ordering of these choices.)

(Cutler & Carter, 1987, S.139)

Die rhythmische Segmentierungshypothese umfasst also zwei verschiedene Komponenten: eine rein datengetriebene Komponente, die starke Silben im sprachlichen Signal markiert. Die so markierten starken Silben dienen dann zur Initiierung von lexikalischen Suchprozessen im Hauptlexikon. Dabei wird zunächst angenommen, dass alle der starken Silben folgenden Silben zu dem lexikalischen Eintrag gehören. Die zweite Komponente nutzt nun lexikalische Information: stösst der Suchprozess in der Eingabekette auf eine unbetonte Silbe, die zusammen mit den vorherigen Silben nicht mehr in eine Wortform integriert werden kann, zu der ein lexikalischer Eintrag gefunden werden kann, muss diese Silbe der Anfang eines weiteren Wortes sein.

Um die Effizienz der datengetriebenen Komponente dieser Strategie für das Englische zu überprüfen, führten Cutler und Carter (1987) eine Analyse der Häufigkeit verschiedener Wortbetonungsmuster in einem Korpus gesprochener englischer Sprache mit ca. 180000 Worttokens durch. Sie fanden, dass ca. 90% der Inhaltswörter dieses Korpus tatsächlich initial eine starke Silbe (Primär- oder Sekundärakzent) aufwiesen, wohingegen dies - unter der Annahme, dass einsilbige Funktionswörter als schwache Silbe realisiert werden - nur bei ca. 10% der Funktionswörter zutrif. Nach diesen

Zahlen ist zumindest für die Identifizierung der Anfänge von Inhaltswörtern im Englischen eine metrische Segmentierungsstrategie recht erfolgreich. Mit dieser Strategie sollten 85% aller Segmentierungspunkte auch gleichzeitig Wortanfänge sein, in 15% würde der gefundene Segmentierungspunkt nicht mit einem Wortanfang korrelieren und somit zu einer Fehlsegmentierung führen. Problematisch bleibt jedoch, wie anhand dieser Segmentierungsstrategie Funktionswörter effizient erkannt werden können, da anhand dieser Strategie nur die Anfänge initial betonter Funktionswörter direkt identifiziert werden können. Wie die obigen Prozentzahlen zeigen, sind jedoch nur ein geringer Anteil der Funktionswörter initial betont. Nach dem Ansatz von Cutler und Carter (1987) können unbetonte Funktionswörter erst postlexikalisch nach dem Zugriff auf die Inhaltswörter als eigenständige Elemente erkannt werden, nämlich dann, wenn klar wird, dass die entsprechende unbetonte Silbe nicht mehr zu dem sich gerade in der Verarbeitung befindlichen Inhaltswort gehören kann. Die Tatsache, dass Funktionswörter jedoch im allgemeinen auch im Satzkontext viel schneller erkannt werden können als Inhaltswörter (Friederici, 1983; Friederici, 1985), lässt jedoch Zweifel an der Annahme aufkommen, dass ihre Verfügbarkeit abhängig von der vorherigen Identifizierung vorangehender Inhaltswörter ist.

Neben den rein sprachstatistischen Argumenten für die Effizienz einer rhythmusbasierten Segmentierungsstrategie zumindest für die Erkennung von Inhaltswörtern im Englischen gibt es auch empirische Evidenz dafür, dass Sprecher des Englischen starke Silben als potentielle Wortanfänge heranziehen. In einer sogenannten *Word-Spotting* Aufgabe, in der die Probanden entscheiden sollten, ob ein auditiv präsentierte mehrsilbiges Pseudowort ein existierendes Wort enthält oder nicht, zeigten sich deutliche Effekte der rhythmischen Struktur dieser Pseudowörter. Wenn ein eigentlich einsilbiges Wort im Pseudowort durch eine Silbengrenze durchschnitten wurde (z. B. MIN-Tef) zeigten sich nur dann größere Schwierigkeiten bei der Erkennung dieses Wortes gegenüber eingebetteten Wörtern, die einer Silbe des Pseudowortes entsprachen (z.B. MINT-kef), wenn die zweite Silbe eine starke Silbe war, d. h. einen Vollvokal aufwies, nicht aber wenn diese eine schwache Schwa-Silbe war (Cutler & Norris, 1988). Im Einklang mit den vorher beschriebenen einzelsprachspezifischen Segmentierungsstrategien fand sich bei einer analogen Untersuchung mit Sprechern des Französischen kein Effekt des metrischen Status der

zweiten Silbe auf die Reaktionszeiten (Cutler et al., 1992). Auch bei sogenannten "Verhörern", bei denen der Hörer eine Wortgrenze an einer Position interpretiert, an der sie vom Sprecher nicht intendiert war, lässt sich der Einfluss rhythmischer Information feststellen. So fanden Cutler und Butterfield (1992) sowohl in spontanen als auch in elizitierten Verhörern von englischen Muttersprachlern, dass Pausen häufiger fälschlicherweise vor starken Silben eingeschoben werden als vor schwachen und dass Pausen häufiger vor schwachen Silben überhört werden als vor starken.

Trotz dieser Evidenzen ist bislang nicht eindeutig belegt, welche Rolle die rhythmische Segmentierungsstrategie innerhalb der lexikalischen Verarbeitungsprozesse erwachsener Sprecher tatsächlich einnimmt. Die experimentellen Befunde lassen nur den Schluss zu, dass englische Sprecher über eine solche Segmentierungsstrategie verfügen und diese in bestimmten metasprachlichen Aufgaben bzw. bei durch beispielsweise Störgeräusche ausgelösten Verarbeitungsproblemen einsetzen. Wie stark dieser Segmentierungsmechanismus zur „normalen, tagtäglichen“ Sprachverarbeitung beiträgt, ist bislang unklar. Dass zur Wortsegmentierung auch lexikalische Information herangezogen wird, scheint deutlich. Zum einen zeigten sich deutliche non-prosodische lexikalische Einflüsse in *Word-Spotting* Aufgaben (McQueen, Norris & Cutler, 1994; Norris, McQueen & Cutler, 1995): so wurden einsilbige Wörter, die in die Anfangssequenz eines anderen Wortes eingebettet waren (z.B. SAC in SACrif) schlechter erkannt, als wenn die Sequenz nicht zu einem Wort zu ergänzen war (z.B. SAC in SACran). Da beide Sequenzen rhythmisch gleich sind, ist zu vermuten, dass der lexikalische Status für diesen Unterschied entscheidend ist: in der Sequenz *sacran* wird spätestens mit dem Auftauchen des Segments /n/ deutlich, dass es sich hierbei nicht um eine Worteinheit handeln kann, da kein mit dieser Sequenz kompatibler lexikalischer Eintrag mehr gefunden werden kann. Anders sind die Verhältnisse in der Sequenz *sacrif*. Hier bleibt bis zum Ende der Sequenz der Lexikoneintrag *sacrifice* als potentieller Kandidat für den Stimulus aktiv und verhindert dadurch zunächst die Segmentierung der Phonemkette nach der initialen Silbe. Dies erschwert offensichtlich die Erkennbarkeit des eingebetteten Wortes.

Gegen die Annahme, dass eine prälexikalische, ausschließlich an prosodischer Information ausgerichtete Segmentierungsstrategie zu den obligatorischen Verarbeitungsmechanismen gehört, spricht auch, dass sich ihre Anwendung unter

experimentellen Bedingungen manipulieren lässt. Sebastián-Gallés et al. (1992) fanden in ihren Untersuchungen mit spanischen und katalanischen Sprechern Veränderungen in den Segmentierungsstrategien in Abhängigkeit von bestimmten Manipulationen der Aufgabenstellungen. Insbesondere bei Verzögerungen der Reaktionen durch das Stellen einer Nebenaufgabe traten Silbeneffekte stärker hervor als bei kürzeren Reaktionszeiten bei Durchführung derselben Aufgabe ohne Nebenaufgabe. Gerade die Tatsache, dass Silbeneffekte in langsameren Reaktionen klarer zu beobachten sind, deutet auf eine metasprachliche Komponente beim Entstehen dieser Effekte hin.

3.1.2 Inputinformation für die rhythmische Segmentierung

Wie oben beschrieben gehen die erwähnten Ansätze davon aus, dass die verschiedenen Segmentierungsstrategien sich anhand bestimmter Inputinformation ausbilden und sie in der Sprachverarbeitung an bestimmter Inputinformation ansetzen. Die Frage, die sich natürlich stellt ist die, welche Inputinformation dies sein kann, d.h. gibt es phonetisch eindeutige Korrelate.

3.1.2.1 Phonetische Korrelate der Wortbetonung

Cutler selbst hat mehrfach die Frage gestellt, welches phonetische Korrelat die Grundlage für die Identifizierung von perzeptuell hervorgehobenen Silben als potentiell Segmentierungspunkt darstellt. Als betonte Silben werden in der Phonetik/Phonologie im allgemeinen solche bezeichnet, die durch einen oder mehrere phonetische Parameter gegenüber ihrer Umgebung perzeptuell hervorgehoben sind. Zu diesen phonetischen Parametern werden primär die Silbendauer, die Tonhöhe, sowie die Lautstärke gezählt, wobei zwischensprachliche Unterschiede in der Verwendung dieser einzelnen Parameter zu beobachten sind. So gibt es Sprachen, die eher die Tonhöhe einsetzen, andere, die eher die Variation in der Silbendauer einsetzen, in anderen korrelieren die Merkmale stark. Beispielsweise ist für das Deutsche die Vokaldauer eines der stabilsten Korrelate von Wortakzent (Jessen, Marasek, Schneider & Clahßen, 1995), während im Französischen diesbezüglich eher die Tonhöhe relevant ist

(Meisenburg & Selig, 1998). Das Englische nutzt als suprasegmentale phonetische Merkmale für Betonung sowohl die Vokaldauer als auch die Tonhöhe (Cutler, Dahan & van Donselaar, 1997). Für den englischen Hörer scheinen jedoch weniger diese suprasegmentalen Merkmale als die Vokalqualität die entscheidende Information für die Wahrnehmung rhythmischer Eigenschaften zu sein. Im Englischen unterscheiden sich betonte und unbetonte Silben in den meisten Fällen neben den genannten Merkmalen auch dadurch, dass betonte Silben einen Vollvokal aufweisen, während in unbetonte Silben ein Reduktionsvokal, im allgemeinen ein Schwa, auftritt.

Cutler (1993) plädiert zunächst aus rein theoretischen Überlegungen für eine höhere Relevanz der Vokalqualität gegenüber suprasegmentalen Eigenschaften für die Sprachverarbeitung: Suprasegmentale Korrelate der Betonung sind relationale Eigenschaften, d. h. die perzeptuelle Salienz einer Silbe durch Dauer und Tonhöhe ergibt sich nur dadurch, dass sie in diesen Parametern höhere Werte aufweist als adjazente Silben. Daraus folgt, dass der Betonungsgrad einer Silbe nur unter Berücksichtigung einer größeren Domäne innerhalb des Signals ermittelt werden kann. Nach Cutlers Auffassung ließe sich diese Beschränkung jedoch kaum mit der raschen und effizienten *on-line* Verarbeitung, durch die sich gerade die lexikalische Verarbeitung auszeichnet, vereinbaren. Die Vokalqualität ist demgegenüber eine lokale Eigenschaft, die an einem einzelnen Segment festgemacht werden kann und somit dem System eine sehr schnelle Identifizierung des Betonungsmusters erlaubt. Experimente mit englischen Sprechern zeigen denn auch, dass diese stärker auf die Betonungsverschiebung bei zweisilbigen Wörtern mit einer Schwa-Silbe reagieren als bei Wörtern, in denen beide Silben einen Vollvokal haben (Cutler & Clifton, 1984). Für eine relativ geringe Rolle von suprasegmentaler Information für die lexikalischen Zugriffsprozesse sprechen zudem Befunde, nach denen Homophone, die sich ausschließlich durch das Betonungsmuster unterscheiden, lexikalische Aktivierungsprozesse für beide Homophone auslösen (Cutler, 1986).

3.1.2.2 Phonetische Korrelate von Rhythmustypen

Wenn sich – wie von Cutler und Mehler angenommen – der Segmentierungsmechanismus während des frühen Spracherwerbs den rhythmischen

Gegebenheiten der Umgebungssprache anpasst, so ist zu fragen, welche Information im sprachlichen Signal mit den angenommenen Rhythmustypen korreliert. Die Grundlage der typologischen Unterscheidung verschiedener Rhythmustypen wie sie von Abercrombie (1967) vorgeschlagen wurde, nämlich die Annahme, dass sich die Sprachen in ihrer rhythmischen Grundorganisation unterscheiden, ließ sich bislang phonetisch kaum nachweisen. Wenn in silbenzählenden Sprachen die Silbe die rhythmische Grundeinheit darstellt, so folgt daraus, dass die einzelnen Silben in zeitgleichen Abständen aufeinander folgen und somit ihre Länge nicht sehr variabel sein kann. Demgegenüber sollten in akzentzählenden Sprachen die betonten Silben in zeitgleichen Abständen auftreten. Da zwischen den betonten Silben eine variable Menge unbetonter Silben auftreten kann, folgt daraus auch, dass in akzentzählenden Sprachen die Silbendauer weitaus variabler sein sollte als in silbenzählenden Sprachen. Phonetische Analysen brachten jedoch wenig Evidenz für diese sogenannte Isochronie-Hypothese. So ist beispielsweise im Englischen entgegen den Voraussagen der Isochronie-Hypothese der zeitliche Abstand zwischen den betonten Silben abhängig von der Anzahl und der Komplexität der Silben, die ein Fuß enthält, und von der seriellen Position des Fußes in der gesamten Äußerung (Bolinger, 1965; Dauer, 1983). In den silbenzählenden Sprachen Spanisch und Französisch erwies sich die Silbendauer ebenfalls als variabel (Borzone de Manrique & Signorini, 1983; Dauer, 1983; Wenk & Wiolland, 1982). Vergleiche der Dauerverhältnisse in silben- und akzentzählenden Sprachen zeigten weder systematische Unterschiede in der Variabilität der Silbendauer noch systematische Unterschiede in der Variabilität der Abstände zwischen betonten Silben zwischen den Sprachen (Dauer, 1983; Roach, 1982). Es wurde daher vorgeschlagen, dass der Eindruck, Sprachen würden unterschiedlichen Rhythmustypen angehören, nicht direkt auf Differenzen in der zeitlichen Strukturierung beruht, sondern ein Produkt verschiedener phonologischer Merkmale ist (Auer & Uhmann, 1988; Dauer, 1983). Dabei scheint die Zuordnung einer Sprache zu einem Rhythmustyp mit den folgenden Merkmalen zu korrelieren:

- akzentzählende Sprachen erlauben komplexere Silbenstrukturen und damit auch schwerere Silben als silbenzählende Sprachen

- in akzentzählenden Sprachen ist die Akzentzuweisung abhängig vom Silbengewicht, in silbenzählenden Sprachen spielt das Silbengewicht keine Rolle für die Akzentzuweisung
- in akzentzählenden Sprachen ist die Position des Wortakzents variabel, Wortakzent kann distinktiv eingesetzt werden, in silbenzählenden Sprachen ist der Wortakzent fest und wird nicht distinktiv eingesetzt
- akzentzählende Sprachen haben ein System von Reduktionsvokalen, die in unbetonten Silben auftauchen, silbenzählende Sprachen haben dagegen ein stabiles Vokalsystem und Vokalharmonie ist möglich
- akzentzählende Sprachen haben schlecht definierte Silbengrenzen mit ambisilbischen Konsonanten und Geminatenreduktion, während silbenzählende Sprachen klare Silbengrenzen aufweisen

Nun kann jede Sprache die oben formulierten Merkmale in verschiedenen Ausprägungskombinationen aufweisen. So hat beispielsweise das Spanische wie das Englische variable Wortakzentpositionen und kann Akzent distinktiv einsetzen, es hat jedoch im Gegensatz zum Englischen klare Silbengrenzen und somit keine ambisilbischen Konsonanten sowie keine Vokalreduktion. Das Katalanische gleicht dem Englischen mit variabler Akzentposition und Vokalreduktion in unbetonten Silben, es setzt Wortakzent jedoch nicht kontrastiv ein, hat klare Silbengrenzen und keine ambisilbischen Konsonanten (Sebastián-Gallés et al., 1992). Diese beiden Beispiele zeigen, dass die oben genannten phonologischen Merkmale in verschiedenen Sprachen in unterschiedlicher Bündelung vorkommen. Damit wird einer typologischen Klassifizierung zwischen akzent- und silbenzählend als sich ausschließende Kategorien die Grundlage entzogen, vielmehr wird die Differenzierung zwischen akzent- und silbenzählenden Sprachen zu einem Kontinuum, auf dem jede Sprache – je nachdem wieviele Merkmale der einen oder der andere Kategorie sie aufweist - ihren spezifischen Platz einnimmt.

Es fragt sich jedoch, inwieweit diese Merkmale tatsächlich geeignet sind, Sprachen, die anhand ihres rhythmischen Eindrucks verschiedenen Kategorien zugeordnet wurden, in analoger Weise zu kategorisieren. Ramus, Nespór und Mehler (1999) gingen dieser Frage für eines der oben genannten phonologischen Merkmale,

nämlich der Komplexität von Silbenstrukturen, im Detail nach. Ihrer Analyse liegt die folgende Annahme zugrunde: wenn ein wesentliches Korrelat der Wahrnehmung verschiedener Rhythmustypen im Grad der Silbenkomplexität einer Sprache besteht, so sollten sich zwischensprachliche Unterschiede in Unterschieden im Vokal-Konsonant-Verhältnis ausdrücken. Die einfachste universale Silbenstruktur ist die CV-Struktur. Wenn also eine Sprache ausschließlich diesen Silbentyp zulässt, sollte das Verhältnis von Vokalen und Konsonanten in den Wörtern dieser Sprache relativ ausgewogen sein. Komplexere Silbenstrukturen zeichnen sich lediglich durch einen höheren Anteil von Konsonanten gegenüber einfachen CV-Silben aus, der vokalische Anteil bleibt von Unterschieden in der Silbenkomplexität unberührt. Aus diesem Grund ist zu erwarten, dass in Sprachen, die relativ komplexe Silben zulassen, deutlich mehr Konsonanten als Vokale in sprachlichen Äußerungen zu finden sind. Dies überprüften die Autoren an Korpora der Sprachen Englisch, Niederländisch, Polnisch, Französisch, Spanisch, Italienisch, Katalanisch und Japanisch, womit Vertreter aller traditionellen Rhythmustypen in die Analyse einbezogen waren. In Bezug auf zwei Analysemerkmale bildeten die Sprachen Gruppen, die mit den traditionellen Rhythmustypen korrelieren. So fand sich bei den traditionell zum akzentzählenden Typus gerechneten Sprachen Englisch, Niederländisch und Polnisch ein relativ geringer Vokalanteil bei hoher Variabilität der Dauer der konsonantischen Anteile. Die traditionell silbenzählenden Sprachen Spanisch, Italienisch, Französisch und Katalanisch wiesen dagegen übereinstimmend einen höheren Vokalanteil bei einer niedrigeren Variabilität der konsonantischen Anteile auf als die akzentzählenden Sprachen. Das Japanische lag in den Ergebnissen dieser Analyse sehr weit von den anderen Sprachen entfernt: es zeigte bei weitem den höchsten Vokalanteil bei geringster Variabilität der Dauer der konsonantischen Anteile.

Die Ergebnisse dieser Analyse deuten also an, dass die Anteile von Vokalen und Konsonanten im Signal mit der Wahrnehmung verschiedener rhythmischer Typen korreliert. Weitere Analysen dieser Art wären sicherlich notwendig, um auch die Relevanz der übrigen Merkmale zu überprüfen. Geht man davon aus, dass die Differenzierung zwischen Vokalen und Konsonanten tatsächlich bereits auf der Wahrnehmungsebene eine sehr grundlegende Kategorisierung darstellt, könnte dies tatsächlich einen Typ von Information darstellen, der bereits von Säuglingen genutzt

wird, um eine Segmentierungsstrategie auszubilden. Aber hier sind sicherlich weitere Forschungsarbeiten vonnöten.

3.2 Syntaktische Verarbeitung

3.2.1 Prosodische Hinweise auf grammatische Klassen

Diskutiert wurde auch die Möglichkeit, prosodische Merkmale frühzeitig im Verarbeitungsprozess zur syntaktischen Klassifizierung von lexikalischen Einheiten heranzuziehen (als Überblick s. Kelly, 1996). Diese Diskussion konzentriert sich auf zwei grundlegende Klassifikationen: zum einen eine mögliche phonologische Differenzierbarkeit von Elementen der geschlossenen und Elementen der offenen Klasse und zum anderen auf phonologische Unterschiede zwischen Nomen und Verben, wobei bisherige Forschungsarbeiten fast ausschließlich zum Englischen durchgeführt wurden.

Im Englischen besteht der zentrale prosodische Unterschied zwischen Elementen der offenen und Elementen der geschlossenen Klasse darin, dass erstere mindestens eine prosodisch starke Silbe, d. h. eine betonte Silbe bzw. eine Vollvokalsilbe aufweisen, während Elemente der geschlossenen Klasse in kontinuierlicher Sprache auch ohne starke Silbe realisiert werden können. Dies ist tatsächlich auch die typischere Realisierung dieser Elemente. So fanden Cutler und Carter (1987) in ihrer bereits erwähnten Korpusanalyse gesprochenen Englischs, dass nur 14% aller starken Silben in einem Element der geschlossenen Klasse auftraten, aber 72% aller schwachen Silben zu einem Element der geschlossenen Klasse gehörten. Die Evidenz, dass diese Korrelation für die Verarbeitung gesprochener Sprache tatsächlich eine Rolle spielt, ist allerdings dünn und eher indirekt. Cutler (1993) berichtet über Befunde, nach denen englische Muttersprachler die Grammatik einer künstlichen Sprache besser lernen, wenn deren prosodische Verhältnisse denen des Englischen entsprechen, d. h. wenn Elemente der geschlossenen Klasse als schwache Silben realisiert werden und Elemente der offenen Klasse zumindest eine starke Silbe involvieren. Auch die Ergebnisse der schon zitierten

Verhöreranalyse von Cutler und Butterfield (1992) deuten in diese Richtung: in dem Corpus von Fehlsegmentationen des sprachlichen Inputs zeigte sich, dass starke Silben häufiger vom Hörer als ein Element der offenen Klasse interpretiert werden, während schwache Silben häufiger als ein Element der geschlossenen Klasse identifiziert werden. So wird beispielsweise beim Missverstehen der Äußerung *economist* als *the communist* die schwache Silbe als Funktionswort *the* interpretiert, die starke Silbe *co* bleibt Teil eines Elements der offenen Klasse.

Die Ergebnisse einer akustischen Analyse der phonetischen Eigenschaften der Vokale von Inhaltswörtern und Funktionswörtern lassen jedoch Zweifel daran aufkommen, dass die beiden Klassen klar allein anhand rhythmisch-prosodischer Merkmale differenzierbar sind. Morgan, Shi und Allopenna (1996) analysierten für Inhalts- und Funktionswörter aus einem Korpus kindgerichteter Sprache Vokallänge, Silbenamplitude, Grundfrequenzveränderungen und Vokalqualität als mögliche Korrelate der prosodischen Stärke eines Elements. Zwar fanden die Forscher signifikante durchschnittliche Unterschiede zwischen Inhalts- und Funktionswörtern in allen der gemessenen Parameter, aber die Verteilung der einzelnen Parameter für die beiden Wortklassen überschneidet sich stark, so dass die Definition eines *cut-off points*, der zwischen den beiden Wortklassen trennt, für keinen der einzelnen Parameter möglich war. Trotzdem konnten die Autoren zeigen, dass unter gleichzeitiger Berücksichtigung dieser akustischen Parameter und einer Reihe anderer Merkmale wie Silbenzahl, Silbenkomplexität, Frequenz und Äußerungsposition eine korrekte Kategoriezuweisung der Inhalts- und Funktionswörter in 80 bis 90% der Fälle gelingt. Sie plädieren deshalb für ein Modell, dass parallel verschiedene Typen von Hinweisen im sprachlichen Input für eine grobe Kategorisierung dieser Wortklassen evaluieren und kombinieren kann.

Die Frage nach Möglichkeiten einer feineren Differenzierbarkeit von Wortklassen innerhalb der Elemente der offenen Klasse ist ebenfalls vor allen Dingen im Englischen für Nomen und Verben untersucht worden (Kelly, 1996). Im Englischen könnte das Betonungsmuster als Hinweis auf die Klassenzugehörigkeit eines lexikalischen Elements herangezogen werden, da in der Klasse der Nomen das trochäische Betonungsmuster stärker dominiert als in der Klasse der Verben. Dies zeigt sich besonders deutlich in der Klasse von homographen Formen, die sowohl als Nomen

als auch als Verben gebraucht werden können (z.B. *permit*). Wenn diese Formen zwischen den Wortklassen im Betonungsmuster variieren, weist die Nomenform stets ein trochäisches und die Verbform stets ein jambisches Betonungsmuster auf. Zudem weisen zweisilbige englische Verben grundsätzlich häufiger ein jambisches Betonungsmuster auf als zweisilbige Nomen (Kelly & Bock, 1988; Kelly, 1992). Tatsächlich scheinen sich Sprecher des Englischen dieser Korrelation von Betonung und Wortklassenzugehörigkeit auch bewusst zu sein. Bei der Aufgabe, Pseudowörter in einem Satz laut zu lesen, in dem sie durch den syntaktischen Kontext entweder als Nomen oder als Verben festgelegt waren, wurden die Pseudowörter in einem Nomenkontext häufiger auf der ersten Silbe betont produziert als in einem Verbkontext (Kelly & Bock, 1988). Bei der Aufgabe, auditiv präsentierte zweisilbige Wörter als Nomen oder Verben zu klassifizieren, zeigten englische Muttersprachler eine Interaktion zwischen Betonungsmuster und Schnelligkeit der Kategorisierung: Trochäische Nomen wurden schneller korrekt klassifiziert als trochäische Verben während jambische Verben schneller korrekt klassifiziert wurden als jambische Nomen (Kelly, 1996).

Aufgrund der Ergebnisse einer Computersimulation kamen Durieux und Gillis (2001) allerdings zu dem Schluss, dass bei alleiniger Berücksichtigung prosodischer Eigenschaften die Prädiktion der Wortklasse von zweisilbigen Nomen und Verben nur bedingt möglich ist. Auf Grundlage der Vorgabe eines Korpus von zweisilbigen Nomen und Verben, bei denen Betonungsmuster und Wortklasse kodiert war, war das System in der Lage, 69% von „neuen“ zweisilbigen Wörtern, bei denen nur das Betonungsmuster nicht aber die Wortklasse kodiert war, korrekt zu klassifizieren. Allerdings war der Unterschied zwischen korrekt klassifizierten Nomen und korrekt klassifizierten Verben erheblich: während Nomen in 92% der Fälle korrekt der Wortklasse zugewiesen wurden, galt dies nur für 38% der Verben. Dieses Verhältnis deutet auf einen *Nomen-Bias* des Algorithmus hin. Offensichtlich wurden generell mehr Nomen- als Verbentscheidungen getroffen, was die hohe Anzahl korrekter Entscheidungen in der Nomenklasse bedingen könnte. Diese Präferenz für Nomen-Entscheidungen könnte damit zusammenhängen, dass das Trainingskorpus fast doppelt so viele Nomen wie Verben enthielt. Diese Nomendominanz mag zu einer besonders starken Assoziation zwischen dem trochäischen Betonungsmuster und der Wortklasse

Nomen geführt haben, so dass eine Vielzahl der trochäischen Verben des Testkorpus ebenfalls als Nomen klassifiziert wurden. Es ist jedoch anzunehmen, dass in vielen Sprachen auch im natürlichen sprachlichen Input mehr Nomen als Verben zu finden sind, so dass auch das menschliche Verarbeitungssystem bei der Beachtung entsprechender Information zu einer zu häufigen Zuordnung eines Wortes zur Klasse der Nomina kommen würde.

Neben der Frage nach der Zuverlässigkeit prosodischer Merkmale als Klassifizierungsgrundlage stellt sich die Frage nach ihrer tatsächlichen Rolle für die Sprachverarbeitung. Die bisher zitierten Befunde zeigen lediglich, dass der Sprecher/Hörer bestimmte Korrelationen zwischen prosodischen und syntaktischen Merkmalen lexikalischer Elemente für die Lösung bestimmter Aufgaben ausnutzen kann. Dies wurde aber im wesentlichen bei der Lösung metasprachlicher Aufgaben wie z. B. der Wortkategorisierung festgestellt, bzw. im Fall der Verhöreranalyse in Situationen, in der die Äußerung des Sprechers nicht korrekt verarbeitet wurde. Es ist unklar, inwieweit die berichteten Verhörere von Reanalyse- und Interpretationsprozessen geprägt sind, die eben nicht nur stimulus- sondern auch wissensgeleitet im Sinne eines *top-down* Prozesses versuchen, dem akustischen Stimulus eine Struktur bzw. eine Bedeutung zuzuordnen. Bislang liegt keine empirische Evidenz vor, die klar zeigt, dass die Ausnutzung dieser Korrelationen zu den normalen Sprachverarbeitungsroutinen gehört.

3.2.2 Prosodische Phrasierung als Hinweis auf die syntaktische Struktur

Der Frage nach der Rolle prosodischer Information für die Satzverarbeitung ist erst in jüngerer Zeit stärker in das Zentrum psycholinguistischer Forschung gerückt. Lange Zeit orientierten sich Modelle zur Satzverarbeitung an Befunden aus Untersuchungen zum Lesen, d. h. zur Verarbeitung geschriebener Sprache. Dies gilt unter anderem auch für das einflussreiche sogenannte *Gardenpath*-Modell von Frazier (1978, als neuerer Überblick Frazier & Clifton, 1996). Nach diesem modularen Modell geschieht die Berechnung der syntaktischen Struktur eines Satzes nach rein syntaktischen Prinzipien: der Verarbeitungsmechanismus erstellt zunächst aufgrund der Wortkategorieinformation der lexikalischen Einheiten eines Satzes zusammen mit

generellen Prinzipien eine möglichst einfache syntaktische Struktur der Äußerung und zwar fortlaufend während der Verarbeitung des Satzes. Einfachheit definiert sich dabei über zwei grundlegende Verarbeitungsprinzipien, dem sogenannten *Minimal Attachment* Prinzip und dem sogenannten *Late Closure* Prinzip. Grundsätzlich wird dem Input zunächst eine Struktur zugewiesen, die diesen Prinzipien entspricht und erst in dem Fall, in dem der weitergehende Input nicht mehr mit dieser anfänglichen Strukturzuweisung kompatibel ist, wird die Möglichkeit einer anderen syntaktischen Struktur des Inputs geprüft. Diese erneute Analyse des Inputs wird als Reanalyse bezeichnet.

Nach dem *Minimal Attachment* Prinzip soll für jedes neue Wort einer Äußerung eine syntaktische Anbindung an die vorangegangene Struktur geschaffen werden, die zu einer möglichst flachen Struktur führt, d. h. eine Struktur mit geringer Anzahl von syntaktischen Knoten wird gegenüber einer Struktur mit einer höheren Anzahl von Knoten bevorzugt. Nach dem *Late Closure* Prinzip werden neu in die Struktur einzubindende Wörter bevorzugt in die gerade in der Bearbeitung befindliche Konstituente eingefügt und nicht zur Öffnung einer neuen Konstituenten verwendet.

Die Wirkungsweisen dieser Prinzipien sollen an dem folgenden in der Literatur häufig verwendeten Beispiel (z.B. Frazier & Clifton, 1996) deutlich gemacht werden. Dem Satz *The horse raced past the barn fell* können bis zum letzten Wort *fell* zwei alternative Strukturen zugewiesen werden. Bedingt durch die morphologische Mehrdeutigkeit der Verbform *raced* kann dieses als finites Verb eines Matrixsatzes mit dem Subjekt *the horse* fungieren, es könnte aber auch einen Nebensatz einleiten. Im zweiten Fall würde sich zwischen *horse* und *race* eine *Clause*-Grenze befinden. Nach beiden Verarbeitungsprinzipien sollte hier zunächst die zuerst genannte Struktur zugewiesen werden, da nur in diesem Fall die Verbform *raced* in die begonnene Satzkonstituente direkt integriert werden kann und diese Lesart zudem eine flachere syntaktische Struktur aufweist als die zweite, die eine Satzeinbettung erfordert.

Für die Verarbeitungsrelevanz dieser Strategien spricht eine Vielzahl von experimentellen Befunden. Bei der Untersuchung dieser Prinzipien wurde vor allen Dingen mit Sätzen gearbeitet, die lokale syntaktische Ambiguitäten aufweisen, die durch verschiedene Möglichkeiten der Positionierung syntaktischer Grenzen innerhalb der Sätze bedingt sind und im Verlaufe des Satzes aufgelöst werden. Sowohl die

Ergebnisse von behavioralen Leseexperimenten (z. B. Frazier & Rayner, 1982; Ferreira & Clifton, 1986) als auch die Ergebnisse von Studien mit ereigniskorrelierten Potentialen (z.B. Osterhout, Holcomb & Swinney, 1994; Mecklinger, Schriefers, Steinhauer & Friederici, 1995) stützen die Vermutung, dass bei der Verarbeitung struktureller Ambiguitäten tatsächlich zunächst nur eine syntaktische Struktur berechnet wird und bei Auflösung der Ambiguität - falls notwendig - eine Reanalyse stattfindet.

Es stellt sich die Frage, inwieweit diese Prinzipien auch bei der Verarbeitung gesprochener Sprache eine Rolle spielen. Zwar ist die sequentielle Präsentation von Textteilen, wie sie in den Leseexperimenten vorgenommen wurde, der Situation in der auditiven Sprachwahrnehmung vergleichbar, da das akustische Signal eine zeitliche Ausdehnung hat und somit sukzessiv verfügbar und analysierbar wird, aber es besteht ein anderer, in diesem Zusammenhang wesentlicher Unterschied zwischen gesprochener und geschriebener Sprache: in der gesprochenen Sprache sind syntaktisch relevante Grenzen sehr häufig durch prosodische Merkmale markiert. Wenn diese Merkmale direkt in die Analyse einbezogen werden, sollten mögliche syntaktische Ambiguitäten in der gesprochenen Sprache an Ort und Stelle aufgelöst werden können, d. h. der Verarbeitungsmechanismus sollte sofort, wenn im Signal das Fragment *horse raced* erreicht wird, entscheiden können, ob zwischen diesen Wörtern eine *Clause*-Grenze vorliegt oder nicht. Falls dies zutrifft, sollten strukturell bedingte Fehlinterpretationen und Reanalyseprozesse bei lokalen syntaktischen Ambiguitäten in der Verarbeitung gesprochener Sprache nicht zu beobachten sein.

Genau dies ist mittlerweile in einer Reihe von Studien untersucht worden. Eine der ersten Arbeiten stammt von Marslen-Wilson, Tyler, Warren, Grenier und Lee (1992). Sie überprüften die Anwendung der *Minimal Attachment* Strategie bei der Präsentation gesprochener Sprache anhand von Sätzen des folgenden Typs:

- (a) The workers considered the last offer from the management / of the factory
- (b) The workers considered the last offer from the management / was a real insult

Bis zur durch den Schrägstrich gekennzeichneten Stelle sind diese Sätze strukturell ambig. Die Ambiguität beruht darauf, dass das Verb *considered* zwei verschiedene Argument-Strukturen zulässt, es kann entweder ein direktes Objekt als Ergänzung

aufweisen wie in (a) oder einen Komplementsatz wie in (b). Die eingesprochenen Sätze wurden an der durch den Schrägstrich gekennzeichneten Stelle geschnitten. In einer *cross-modalen* Benennaufgabe wurden den Probanden beide Anfangsfragmente auditiv präsentiert und im direkten Anschluss an jedes Fragment das Wort visuell präsentiert, das dem Fragment der nicht dem *Minimal Attachment* entsprechenden Lesart im Satz gefolgt war, im obigen Beispiel also das Kopulaverb *was*. Die Aufgabe der Probanden bestand darin, dieses Wort laut zu lesen. Diesem experimentellen Design lag folgende Überlegung zugrunde: Wenn die Hörer grundsätzlich eine *Minimal Attachment* Strategie ohne Einbezug prosodischer Markierungen anwenden, sollte die syntaktische Struktur, die dem initialen Satzfragment zugewiesen wird stets die Objekt-Lesart sein. In diesem Fall sollten sich keine Unterschiede in der Lesezeit für *was* in Abhängigkeit vom vorhergehenden Satzfragment zeigen. Wird allerdings die prosodische Information der Satzfragmente in die syntaktische Analyse einbezogen, dann kann das Kopulaverb bei Satztyp (a) nicht in die bis dahin aufgebaute Struktur für das initiale Fragment integriert werden, wohl aber bei Satztyp (b). Dies sollte sich in einer Verlängerung der Lesezeiten für das kritische Wort im Satztyp (a) gegenüber dem Satztyp (b) ausdrücken. Genau dieses Leistungsmuster fand sich auch in den Reaktionszeiten der Probanden.

In ähnlicher Weise gingen Warren, Grabe und Nolan (1995) in einem Experiment vor, das die Anwendung der *Late Closure* Strategie bei der Verarbeitung von Sätzen mit lokalen syntaktischen Ambiguitäten des folgenden Typs überprüfte:

- (a) Whenever parliament discusses Hong Kong // problems (can be solved)
- (b) Whenever parliament discusses Hong Kong problems // (they can be solved)

Im Experiment wurden den Probanden Satzfragmente, die vor der strukturellen Auflösung der syntaktischen Ambiguität endeten, präsentiert. Mit Ende des Fragments wurde ein Wort visuell präsentiert, das die Probanden zu benennen hatten. Dieses Wort war in allen Fällen nur in die *Early Closure* Lesart, d. h. Lesart (a), zu integrieren, d. h. für die Satzfragmente (a) und (b) konnte dieses Wort beispielsweise das Verb *arise* sein. Die Probanden zeigten schnellere Benennzeiten für die Wörter, wenn ihnen ein *Early Closure* Satzfragment (a) voranging. Dieses Muster zeigt, dass den Sätzen schon vor der Verarbeitung des disambiguierenden Wortes die richtige Lesart zugewiesen worden

war. Da sich – wie Warren und Mitarbeiter in einer detaillierten phonetischen Analyse zeigen – die Sätze im Bereich der kritischen Grenzen prosodisch deutlich unterschieden, ist davon auszugehen, dass diese frühe Disambiguierung durch das Vorhandensein bzw. Fehlen prosodischer Grenzmarkierungen im Bereich der Wortsequenz *Hong Kong problems* ermöglicht wird.

Anhand eines anderen Paradigmas zeigen auch eigene Arbeiten zum Deutschen die Sensitivität der Hörer für diese prosodischen Grenzmarkierungen (Höhle, Hofmann & Friederici, 1995). In diesem Experiment wurden lokal ambige Satzpaare des folgenden Typs benutzt:

- (a) Peter verspricht Anna, zu schlafen und das Licht auszumachen.
- (b) Peter verspricht, Anna zu wecken und das Frühstück zu machen.

Syntaktisch unterscheiden sich diese Sätze systematisch durch die unterschiedliche Position der Grenze zwischen Matrixsatz und Infinitivsatz. In Satztyp (a) lag diese stets hinter der zweiten Nominalphrase, die das Objekt des Matrixsatzes bildete. In Satztyp (b) lag diese Grenze bereits nach dem finiten Verb des Matrixsatzes, die zweite Nominalphrase des Gesamtsatzes war hier immer das Objekt des infiniten Verbs. Diese unterschiedlichen syntaktischen Strukturen ergeben sich durch die Art des infiniten Verbs: eindeutige Strukturen des Typs (a) können nur durch die Verwendung von strikt intransitiven infiniten Verben erzeugt werden, während eindeutige Strukturen des Typs (b) nur durch strikt transitive Verben im Infinitivsatz erzeugt werden können. Mit anderen Worten: die potentielle Ambiguität des initialen Satzteils sollte mit dem ersten Verb des Infinitivsatzes aufgelöst werden.

Sollte diese unterschiedliche syntaktische Strukturierung mit verschiedenen prosodischen Markierungen einhergehen, so ist zu erwarten, dass sich in Satztyp (a) eine prosodische Grenze hinter zweiten Nominalphrase befindet, in Satztyp (b) dagegen vor der zweiten Nominalphrase. Für das Stimulusmaterial wurden die Satzpaare wie im obigen Beispiel ausgesprochen. Anschließend wurden sie digitalisiert und direkt innerhalb des Infinitivpartikels *zu* in zwei Satzfragmente geschnitten.¹ Anschließend

¹ Versuche mit verschiedenen Schnittstellen hatten ergeben, dass dies die Position ist, die zum geringsten, wahrnehmbaren Bruch innerhalb des sprachlichen Reizes führte.

wurden die so erzeugten Satzfragmente gekreuzt wieder zusammengefügt, d. h. das erste Fragment aus Satztyp (a) mit dem zweiten Fragment aus Satztyp (b) und umgekehrt. Dieses Vorgehen hatte zur Folge, dass – wenn die initialen Satzfragmente prosodische Markierungen der syntaktischen Struktur enthalten und diese zur Zuweisung einer syntaktischen Struktur herangezogen werden – in den geschnittenen Satzversionen das Verb nicht in die bis dahin aufgebaute syntaktische Struktur integriert werden kann, was zu Verarbeitungsproblemen gegenüber Sätzen, die nicht geschnitten wurden, führen sollte.

Diese Hypothese wurde anhand einer Aufgabe überprüft, in der den Probanden die Experimentalsätze sowie eine Reihe weiterer Kontrollsätze auditiv mit der Aufgabe vorgegeben wurden, auf eine Reaktionstaste zu drücken, sobald sie in einem der Sätze das Wort *und* hörten. Alle Experimentalsätze enthielten dieses Wort direkt nach dem disambiguierenden infiniten Verb. Probleme, dieses Verb zu integrieren, sollten sich – analog zu den Verlängerungen in Lesezeiten – in einer verlängerten Reaktionszeit auf das Zielwort ausdrücken. Dieses Muster zeigte sich in den Daten zumindest teilweise, nämlich klar für die Sätze mit intransitivem infiniten Verb. Hier zeigten sich signifikant längere Reaktionszeiten für die geschnittenen gegenüber den nicht geschnittenen Sätzen. Bei den Sätzen mit transitivem finiten Verb zeigte sich dagegen kein Unterschied in den Reaktionszeiten zwischen geschnittenen und nicht geschnittenen Sätzen. Die Tatsache, dass sich gerade in den geschnittenen Sätzen mit intransitivem infiniten Verb ein Verzögerungseffekt zeigt, spricht dafür, dass das *Late Closure* Prinzip bei der Verarbeitung gesprochener Sprache nicht unbedingt greift. In der geschnittenen Version folgt dem Matrixverb eine prosodische Grenze. Dies führt offensichtlich dazu, dass die folgende NP *Anna* als nicht mehr in diese Konstituente gehörig interpretiert wird, d.h. es wird eine Struktur verfolgt, die Satztyp (b) entspricht. In diese Struktur passt das intransitive infinite Verb jedoch nicht hinein, da dessen Argumentstruktur verletzt wird. Dies äußert sich in den verzögerten Reaktionen.

Nach den Ergebnissen von Studien mit ereigniskorrelierten Potentialen (EKP) wird die relevante prosodische Information offensichtlich direkt in den Analyseprozess einbezogen und nicht erst um eine Entscheidung zwischen den möglichen Lesarten zu treffen. So finden sich direkte neurophysiologische Korrelate der prosodischen

Phrasierung im EKP, die auf eine unmittelbare Verarbeitung von prosodischer Grenzinformation schließen lassen (Steinhauer, Alter & Friederici, 1999).

3.2.3 Inputinformation für die Erkennung prosodischer Grenzen

Im wesentlichen werden drei phonetische Korrelate syntaktischer Grenzen im sprachlichen Signal diskutiert (z.B. Cooper & Sorenson, 1977; Klatt, 1975; 1976; Cooper & Paccia-Cooper, 1980; Price; Ostendorf, Shattuck-Hufnagel & Fong, 1991; Warren et al., 1995). Phonetische Analysen der Bereiche um syntaktische Grenzen zeigen folgende Merkmale prosodischer Grenzmarkierungen:

- vor syntaktischen Hauptgrenzen findet eine Silbenlängung statt
- vor syntaktischen Hauptgrenzen finden sich Grenztöne in Form von Bewegungen der Grundfrequenz
- syntaktische Hauptgrenzen sind bevorzugte Positionen für Pausen

In einer Reihe von Untersuchungen mit global syntaktisch ambigen Sätzen wurde versucht, den relativen Beitrag, den diese einzelnen Parameter und deren Zusammenwirken für die Satzanalyse liefern, näher zu spezifizieren. Bei den globalen syntaktischen Ambiguitäten handelt es sich ebenfalls um Ambiguitäten, die durch unterschiedliche Möglichkeiten der Phrasenstrukturierung entstanden. Im Gegensatz zu den lokalen syntaktischen Ambiguitäten werden die globalen jedoch im Weiterverlauf des Satzes nicht aufgelöst. Als Beispiel für Sätze, die typischerweise in diesen Studien verwendet wurden, sei der folgende genannt:

The old men and women stayed at home.

Die syntaktische Ambiguität dieses Satzes besteht darin, dass *men and women* als eine einzige Phrase betrachtet werden kann, die insgesamt durch das Adjektiv *old* modifiziert wird, die zweite Lesart teilt diese Sequenz in zwei Phrasen, wobei nur die erste Phrase, d.h. *men* vom Adjektiv modifiziert wird.

Ausgehend von einer Studie von Lehiste (1973), die fand, dass vor allen Dingen die Dauerverhältnisse eingesetzt wurden, um die beiden Lesarten voneinander zu unterscheiden, untersuchten Lehiste, Olive und Streeter (1976) inwieweit durch die Manipulation der Dauerverhältnisse in solchen Sätzen die Präferenz der einen oder anderen Lesart beim Hörer beeinflusst werden kann. Die Forscher fanden, dass eine Kürzung bzw. eine Längung der Wörter vor der potentiellen Phrasengrenze bei einer Konstanthaltung der Grundfrequenz, die Hörer in ihrer Entscheidung, welche Lesart der Sprecher intendiert hatte, signifikant beeinflusste. Scott (1982) konnte diese Ergebnisse mit anderem Satzmaterial replizieren. Sie fand darüber hinaus, dass nicht nur die Dauer eines einzelnen Segments vor einer Phrasengrenze zu deren Wahrnehmung beiträgt, sondern auch die Dauer der gesamten Segmente eines Fußes vor einer Phrasengrenze, d. h. die Verteilung der Längung auf mehrere Silben des phrasenfinalen Fußes führt zum gleichen Wahrnehmungsereignis wie die Längung lediglich der phrasenfinalen Silben um denselben Betrag. Hier zeigt sich, dass die phonetischen Korrelate nicht unbedingt lokal auf einer einzelnen Silbe angesiedelt werden müssen, sondern dass diese über größere Einheiten berechnet werden können.

Ein direkter Vergleich der Wirkung der Manipulation dreier suprasegmentaler Variablen, die als prosodische Grenzmarkierungen diskutiert werden, erfolgte in einer Arbeit von Streeter (1978). Als sprachliches Material verwendete sie algebraische Ausdrücke mit alternativen Lesarten, wie zum Beispiel:

- (a) $(A + E) \times 0$
 (b) vs. $A + (E \times 0)$

Diese algebraischen Ausdrücke wurden einmal für jede der beiden möglichen Klammerungen ausgesprochen und einmal ohne jegliche Klammerung in einer „neutralen“ Phrasierung. Als Basissequenz für die prosodischen Manipulationen wurden aus verschiedenen neutralen Phrasierungen die Wörter herausgeschnitten und zusammengefügt. Aus dieser Basissequenz wurden anschließend je Lesart (a) oder (b) verschiedene Versionen hergestellt, die entweder in ihren Dauerverhältnissen der ursprünglichen Lesart entsprachen, oder in ihrer Grundfrequenzkontur, oder in ihren Amplitudenverläufen. Diese manipulierten Versionen wurden Probanden präsentiert,

mit der Aufgabe zu entscheiden, welche Lesart gemeint ist. Es zeigte sich, dass die Amplitudenverhältnisse allein von den Hörern nicht als reliabler Hinweis auf die gemeinte Lesart interpretiert wurden, jedoch die isolierten Dauerverhältnisse und auch die isolierte Grundfrequenzkontur.

Nach diesen Ergebnissen sind es also insbesondere die Dauerverhältnisse innerhalb des Wortes bzw. der Wörter, die direkt vor einer Phrasengrenze liegen, sowie Veränderungen der Grundfrequenz, die als prosodische Hinweise auf eine syntaktische Grenze vom Hörer herangezogen werden. Zusammenfassend lässt sich also sagen, dass die prosodische Struktur einer Äußerung Hinweise auf die syntaktische Struktur enthält. Dies ergibt sich vor allen Dingen durch eine Übereinstimmung zwischen prosodischen Grenzen und syntaktischen Grenzen, die – wie später noch genauer diskutiert werden wird - jedoch nicht in einer ein-eindeutigen Relation zueinander stehen. Offensichtlich ist der erwachsene Hörer in der Lage, diese Information in die syntaktische Analyse einer Äußerung einzubeziehen. Dabei scheint die prosodische Grenzinformation auch sehr direkt und unmittelbar verarbeitet zu werden, so dass sie von vornherein *on-line* in die syntaktische Verarbeitung einbezogen werden kann.

4 Der prosodische Einstieg in den Spracherwerb

Im vorangegangenen Kapitel wurde gezeigt, dass das akustische Signal Merkmale enthält, die mit bestimmten strukturellen Eigenschaften der zugrundeliegenden Äußerung korrelieren. Des Weiteren haben wir gesehen, dass erwachsene Sprecher/Hörer diese akustischen Merkmale nutzen können, um die Grenzen strukturell - vor allem syntaktisch - relevanter Einheiten zu ermitteln, eventuell auch, um die so gewonnenen Einheiten syntaktisch zu klassifizieren. Dies kann jedoch nicht ohne weiteres auf das noch nicht ausgereifte Sprachverarbeitungssystem des Kindes übertragen werden. Um zu klären, ob Kinder diese Merkmale für die Sprachverarbeitung und den Aufbau sprachlichen Wissens nutzen, müssen die folgenden Fragen beantwortet werden:

- ?? Ab welchem Alter sind Säuglinge sensitiv gegenüber den prosodisch-akustischen Merkmalen, die strukturell relevant sind?
- ?? Ab welchem Alter können Säuglinge diese akustischen Merkmale in ihrer strukturierenden Funktion nutzen?
- ?? Inwieweit sind die relevanten akustischen Merkmale tatsächlich in einer rein datengetriebenen Verarbeitung, d. h. ohne Hinzuziehung höheren lexikalischen oder syntaktischen Wissens, zu interpretieren?

Diese grundlegenden Fragen stellen die Kernpunkte der folgenden Abschnitte dar.

4.1 Sensitivität für prosodische Information

4.1.1 Sensitivität für globale prosodische Information

Es ist anzunehmen, dass die rhythmischen Eigenschaften des sprachlichen Signals zu den ersten Eigenschaften von Sprache gehören, die vom Kind wahrgenommen werden können. Es ist ausserdem zu vermuten, dass gerade die prosodischen Eigenschaften von Sprache dem Säugling bereits pränatal während des letzten Drittels der Schwangerschaft zugänglich sind. In diesem Schwangerschaftabschnitt reagiert der Fötus bereits auf externe Geräusche, was darauf schließen lässt, dass das akustische System funktionsfähig ist. Intra-uterine Aufnahmen haben zudem gezeigt, dass primär die von der Mutter geäußerte Sprache, aber auch sprachliche Äußerungen, die in der Nähe der Mutter getätigt werden, innerhalb des Uterus wahrnehmbar sind. Das sprachliche Signal, das der Fötus im Mutterleib wahrnehmen kann, ähnelt allerdings *low-pass* gefilterter Sprache, da – bedingt durch eine Art Filterfunktion der Gebärmutterwand und des Fruchtwassers - das intra-uterine akustische Signal lediglich die Schallwellen der niedrigen Frequenzbereiche enthält (als Überblick s. Lecanuet & Granier-Deferre, 1993). Sprachliche Signale, die nur die unteren Frequenzbereiche enthalten, zeichnen sich dadurch aus, dass zwar einzelne Lautsegmente nicht mehr identifizierbar sind, die Intonationskontur mit Veränderungen der Grundfrequenz, Intensitätsveränderungen und die rhythmische Strukturierung des Signals durch die Pausensetzung und den Wechsel zwischen betonten und unbetonten Silben jedoch erkennbar bleiben. Befunde, nach denen bereits Neugeborene, die Stimme ihrer Mutter gegenüber einer anderen Frauenstimme bevorzugen und sie zudem eine Präferenz für eine *low-pass* gefilterte Version der Stimme ihrer Mutter gegenüber der ungefilterten Stimme zeigen, sprechen denn auch für pränatale Erfahrungen des Säuglings mit Merkmalen der sprachlichen Äußerungen, die von der Mutter während der Schwangerschaft produziert wurden (DeCasper & Fifer, 1980; Fifer & Moon, 1989).

Eine Reihe von Untersuchungsbefunden deuten zudem darauf hin, dass Neugeborene eine spezielle Sensitivität für prosodische Eigenschaften von Sprache haben. Auch dieses könnte ein Ausdruck dessen sein, dass sie speziell mit den

prosodischen Merkmalen sprachlicher Signal pränatale Erfahrungen haben. Diese besondere Sensitivität für prosodische Merkmale äußert sich beispielsweise in der quasi von Geburt an vorhandenen Fähigkeit, verschiedene Sprachen voneinander zu diskriminieren.

In einer ersten Studie wurde diese Fähigkeit Neugeborener von Mehler und Mitarbeitern (Mehler, Jusczyk, Lambertz, Halsted, Bertoncini & Amiel-Tison, 1988) anhand des *High-Amplitude Sucking* Verfahrens untersucht. Sie fanden, dass französische Säuglinge in der Lage sind, Französisch und Russisch zu unterscheiden, nicht aber Englisch von Italienisch. Zwei Monate alte amerikanische Säuglinge konnten dagegen zwischen Englisch und Italienisch unterscheiden, nicht aber zwischen Russisch und Französisch. Das gleiche Datenmuster zeigte sich sowohl bei der Verwendung *low-pass* gefilterter Sprachstimuli als auch bei der Verwendung nicht gefilterter Sprache. Aus diesem Muster schlossen die Autoren, dass Kinder bereits in ihren ersten Lebensmonaten ihre Umgebungssprache von anderen Sprachen unterscheiden können, aber nicht zwischen verschiedenen Fremdsprachen differenzieren können. Dabei berücksichtigen sie offensichtlich im wesentlichen die prosodischen Eigenschaften der Umgebungssprache. Allerdings kommen Mehler, Dupoux, Nazzi & Dehaene-Lambertz (1996) in einer Reanalyse der Daten von Mehler et al. (1988) zu einem anderen Schluss über die Diskriminationsfähigkeit der Säuglinge zwischen verschiedenen Fremdsprachen. In der Originalanalyse waren die Daten der Kinder, an denen die Diskriminationsfähigkeit zwischen Italienisch und Englisch überprüft worden war, getrennt für die beiden Sprachwechsel ausgewertet worden, d. h. die Kinder, die in der Habituerungsphase Italienisch und in der Testphase Englisch gehört hatten, bildeten eine Untersuchungsgruppe und die Kinder, die in der Habituerungsphase Englisch und in der Testphase Italienisch gehört hatten, bildeten eine andere Untersuchungsgruppe. In der Reanalyse wurden die Daten über die gesamte Gruppe ausgewertet. Hier zeigte sich insgesamt ein Effekt des Sprachwechsels, wobei die Richtung keine Rolle spielte. Nach dieser Reanalyse revidieren die Autoren ihre frühere Feststellung, dass Neugeborene nicht in der Lage sind, Fremdsprachen voneinander zu diskriminieren.

Differenzierter wird das Bild der Fähigkeit, verschiedene Sprachen zu diskriminieren, durch Untersuchungsbefunde von Nazzi, Bertoncini und Mehler (1998). Sie konnten zeigen, dass französische Neugeborene zwar zwischen Englisch und

Japanisch diskriminieren können, nicht aber zwischen Englisch und Niederländisch und auch nicht zwischen Spanisch und Italienisch, wobei Stimuli der beiden letztgenannten Sprachen aber sowohl von englischen als auch von niederländischen Sprachreizen unterschieden werden konnten. Dieses Muster entspricht genau der rhythmischen Differenzierung zwischen den Sprachen, was dafür spricht, dass Neugeborene offensichtlich nur Sprachen unterschiedlicher Rhythmustypen voneinander diskriminieren können: das akzentzählende Englisch vom morenzählenden Japanischen sowie die akzentzählenden Sprachen Englisch und Niederländisch von den silbenzählenden Sprachen Spanisch und Italienisch. Dagegen können dem gleichen Rhythmustyp zugehörige Sprachen nicht voneinander unterschieden werden: weder die beiden akzentzählenden Sprachen Niederländisch und Englisch, noch die beiden silbenzählenden Sprachen Spanisch und Italienisch. Da auch in dieser Untersuchung alle sprachlichen Stimuli *low-pass* gefiltert präsentiert worden waren, sprechen die Befunde deutlich dafür, dass die Fähigkeit, prosodische, insbesondere rhythmische Information wahrzunehmen, von Geburt an präsent ist.

Dieses Muster der Diskriminationsfähigkeit verändert sich während der ersten Lebensmonate noch nicht. Weder 2 Monate alte noch 6 Monate alte mit Englisch als Muttersprache aufwachsende Säuglinge zeigen Anzeichen einer Diskrimination zwischen Niederländisch und Englisch (Christophe & Morton, 1998; Jusczyk, Friederici, Wessels; Svenkerud & Jusczyk, 1993). Auch für 5 Monate alte amerikanische Säuglinge fand sich (Nazzi, Jusczyk & Johnson, 2000), dass sie ihre Muttersprache vom morenzählenden Japanischen unterscheiden können, dass sie zwei Fremdsprachen unterschiedlicher Rhythmustypen (Italienisch vs. Japanisch) diskriminieren können, aber nicht zwei Fremdsprachen des gleichen Rhythmustypus (Italienisch vs. Spanisch). Eine sehr frühe Diskriminationsfähigkeit von rhythmisch ähnlichen Sprachen wurde jedoch in Untersuchungen mit katalanischen und spanischen Säuglingen beobachtet (Bosch & Sebastian-Galles, 1997; 2001). Beide Sprachen werden dem silbenzählenden Sprachtypus zugerechnet. Trotzdem differenzierten 4 Monate alte Säuglinge, die entweder aus monolingualen spanischen oder monolingualen katalanischen Elternhäusern stammten, die beiden Sprachen voneinander, auch wenn ihnen die sprachlichen Stimuli *low-pass* gefiltert präsentiert wurden. Die Divergenz zwischen diesen Ergebnissen und den zuvor berichteten mangelnden

Diskriminationsfähigkeiten von englischen Säuglingen für Englisch und Niederländisch sowie Italienisch und Spanisch könnte zum einen darin begründet sein, dass sich Katalanisch und Spanisch in einem spezifischen prosodischen Merkmal unterscheiden, in dem sich die übrigen Untersuchungssprachen nicht unterscheiden. Es ist jedoch auch möglich, dass die besondere sprachliche Situation in Barcelona, wo die Untersuchungen zum Spanischen und Katalanischen durchgeführt wurden, das Ergebnismuster beeinflussen. Barcelona ist eine hochgradig bilinguale Region, so dass wahrscheinlich ist, dass auch Kinder, die in monolingualen Elternhäusern aufwachsen, bereits relativ früh sprachlichen Input aus beiden Sprachen erhalten. Dies könnte dazu führen, dass die beiden Sprachen besser differenziert werden können, als eine vertraute von einer vollkommen unvertrauten Sprache, wie es in den übrigen Untersuchungen eher der Fall war.

Insgesamt zeigt sich also, dass die frühen Diskriminationsleistungen zwischen Sprachen im wesentlichen von rhythmischen Unterschieden zwischen den Sprachen und der Sensitivität, die bereits Neugeborene für diese Information besitzen, getragen werden. In diesem Zusammenhang, ist auch die Frage diskutiert worden, ob sich bereits in den ersten Lebensmonaten Indikatoren dafür finden lassen, dass die Zielsprache bereits einen herausgehobenen Status für den Säugling hat. Wie erwähnt fanden sich bereits bei Neugeborenen Anzeichen dafür in der Studie von Mehler et al. (1988), die sich später jedoch als Artefakte der Art der Datenauswertung herausstellten (Mehler et al., 1996). Allerdings sprechen auch Befunde von Moon, Panneton-Cooper & Fifer (1993) für einen sehr frühen besonderen Stellenwert der Umgebungssprache. Sie fanden bei zwei Tage alten englischen und spanischen Säuglingen eine Präferenz der jeweiligen Muttersprache zuzuhören gegenüber der jeweiligen anderen Sprache, d. h. englische Säuglinge präferierten Englisch gegenüber Spanisch während die spanischen Säuglinge Spanisch gegenüber Englisch präferierten.

Deutliche Hinweise darauf, dass Säuglinge ihre Muttersprache gegenüber einer anderen Sprache präferieren, zeigten sich auch in den Arbeiten von Bosch und Sebastián-Gallés (1997). Sie fanden, dass ihre monolingualen spanischen und katalanischen 4 Monate alten Säuglinge, sich schneller sprachlichen Stimuli ihrer jeweiligen Muttersprache zuwandten: d. h. Säuglinge aus spanischen Elternhäusern reagierten schneller auf spanische als auf katalanische Sprachreize und Säuglinge aus

katalanischem Elternhaus reagierten schneller auf katalanische als auf spanische Sprachreize. Da sich dieses Muster auch bei Verwendung *low-pass* gefilterter Sprache fand, kann man davon ausgehen, dass es ein prosodisches Merkmal ist, anhand dessen die Kinder ihre Muttersprache von der Nicht-Muttersprache differenzieren.

Es bleibt jedoch noch unklar, aufgrund welcher spezifischen prosodischen Merkmale die Säuglinge in der Lage sind, Sprachen zu unterscheiden. Wie in Abschnitt 3.1.2.2 erläutert, kann nicht davon ausgegangen werden, dass deutliche Unterschiede in den zeitlichen Verhältnissen eine Unterscheidbarkeit der Sprachen der verschiedenen Rhythmustypen zulassen. Es ist fraglich, ob die von Ramus et al. (1999) gezeigte Unterscheidbarkeit akzent-, silben- und morenzählender Sprachen anhand des Grads konsonantischer und vokalischer Anteile als Grundlage für die Differenzierung der Sprachen durch Säuglinge tauglich ist, denn dafür wäre eine wesentliche Voraussetzung, dass bereits Neugeborene in der Lage sind, Vokale von Konsonanten zu unterscheiden und die Auftretenshäufigkeit dieser Klassen in den Stimuli zu bestimmen. Die von Ramus und Mitarbeitern (1999) selbst zitierten Belege für diese Fähigkeit, lassen nur sehr indirekte Schlüsse zu. Sie verweisen zum einen auf Befunde, nach denen Säuglinge auf die Veränderung eines Vokals im sprachlichen Input stärker reagieren als auf die Veränderung eines Konsonanten (Bertoncini, Bijeljac-Babic, Jusczyk, Kennedy & Mehler, 1988). Zum anderen sind Säuglinge von Geburt an für die Anzahl von Silben und damit möglicherweise auch für die Anzahl der Vokale in einem Wort sensitiv (Bertoncini & Mehler, 1981, Bertoncini, Floccia, Nazzi & Mehler, 1995; Bijeljac-Babic, Bertoncini & Mehler, 1993; van Ooyen, Bertoncini, Sansavini & Mehler, 1997). Direkte Evidenz dafür, dass Neugeborene Vokale und Konsonanten verschiedenen Lautklassen zuordnen, fehlt jedoch bislang vollkommen. Es ist außerdem fraglich, wieso die Säuglinge auch bei *low-pass* gefilterte Sprache zu einer Diskriminierung zwischen den Sprachen in der Lage sind, wenn segmentale Eigenschaften für die Fähigkeit grundlegend sein sollen. Das Auftreten eines bestimmten Reaktionsmusters bei *low-pass* gefilterter Sprache wird stets als Beleg dafür angenommen, dass das Reaktionsmuster auf prosodischer Information des sprachlichen Signals beruht und nicht auf segmentalen Eigenschaften. Wenn dies zutrifft, dann kann die Differenzierungsfähigkeit der Säuglinge nicht auf segmentalen Merkmalen beruhen,

sondern muss in Merkmalen gesucht werden, die auch in *low-pass* gefilterter Sprache wahrgenommen werden können.

4.1.2 Sensitivität für spezifische prosodische Merkmale

4.1.2.1 Silbenzahl

Bijeljic-Babic et al. (1993) fanden, dass 4 Tage alte Neugeborene in der Lage sind, die Veränderung der Silbenzahl in sprachlichen Stimuli zu bemerken. Sie präsentierten ihren Probanden dazu in einer Untersuchung mit dem *High-Amplitude Sucking* Paradigma in der Habitierungsphase CV-Silben-Sequenzen mit konstanter Silbenzahl. Nach Erreichen des Habitierungskriteriums wurden der Experimentalgruppe Silbensequenzen mit erhöhter oder reduzierter Silbenzahl geboten. Die Kinder reagierten auf die Veränderung der Silbenzahl mit einem Anstieg der Saugrate. Diese Reaktion zeigte sich jedoch nicht, wenn ihnen nach der Habitierung Stimuli geboten wurden, die dieselbe Silbenzahl aber eine andere Phonemzahl aufwiesen als die Stimuli der Habitierungsphase. Nach diesen Befunden scheinen Neugeborene bereits sensitiv für die rhythmisch relevante Einheit der Silbe. Offensichtlich sind sie jedoch noch nicht in der Lage, subsilbische Einheiten in Form einzelner Laute aus ihrem Input herauszusegmentieren.

Für diesen Schluss sprechen auch weitere Befunde von Bertoncini und Mitarbeitern (Bertoncini et al., 1995). Unter der Hypothese einer anfänglichen universalen Sensitivität gegenüber den rhythmischen Grundeinheiten aller Sprachen verglichen sie, ob französische Neugeborene in gleicher Weise auf die Veränderung der Silbenzahl von japanischen Wörtern reagieren wie auf die Veränderung der Morenzahl. In Bezug auf die Silbenzahl brachte diese Untersuchung die gleichen Ergebnisse wie die von Bijeljic-Babic et al. (1993): die ca. 5 Tage alten Kinder konnten zweisilbige von dreisilbigen Wörtern unterscheiden. Die Kinder diskriminierten allerdings nicht zweisilbige bimoraische Wörter von zweisilbigen trimoraischen. Diese Befunden deuten an, dass nicht alle rhythmischen Einheiten in gleicher Weise der Perzeption zugänglich sind. Die Autoren vermuten, dass Säuglinge zunächst universell silbenähnliche

Repräsentationseinheiten bilden, die erst durch verstärkten sprachlichen Input mit seinen spezifischen rhythmischen Eigenschaften auf die einzelsprachspezifische Einheit adaptiert werden. Daraus würde folgen, dass auch japanische Neugeborene zunächst keine Sensitivität für Moren sondern – wie ihre französischen Altersgenossen – zunächst nur für Silben zeigen. Eine empirische Überprüfung fehlt jedoch bislang.

4.1.2.2 Betonungsmuster

Untersuchungen, die der Frage nach Diskriminationsfähigkeiten von Säuglingen für suprasegmentale Kontraste in Form von Betonungsmustern nachgingen, gehören zu den frühesten Studien im Bereich der frühkindlichen Sprachwahrnehmung. Übereinstimmend konnten Spring und Dale (1977) sowie Jusczyk und Thompson (1978) anhand von Untersuchungen mit dem *High-Amplitude Sucking* Verfahren zeigen, dass Säuglinge bereits in ihren ersten Lebensmonaten zwischen trochäischen und jambischen Zweisilblern unterscheiden können und zwar auch dann, wenn der Betonungsunterschied ausschließlich von extrem kurzen Dauerunterschieden (40 ms) signalisiert wird.

Sansavini, Bertoncini und Giovanelli (1997) bestätigten diese Befunde für Neugeborene. Sie fanden darüber hinaus, dass Neugeborene auch in der Lage sind, Betonungsunterschiede in mehrsilbigen Sequenzen wahrzunehmen (z.B. **TA**-ka-la vs. ta-**KA**-la) (s.a. Sansavini, Bertoncini & Giovanelli, 1994). Im Gegensatz zu vorherigen Untersuchungen, in denen das Stimulusmaterial stets aus nur zwei rhythmisch verschiedenen Zweisilblern mit gleichen Phonemaufbau bestanden hatte (z.B. **BA**ba vs. ba**BA**), präsentierten sie in einem weiteren Experiment eine Vielzahl segmental verschiedener Zweisilbler, bei denen lediglich das Betonungsmuster konstant war (**DA**ga, **NA**ta, **GA**ba. vs. da**GA**, na**TA**, ga**BA**). Auch unter diesen Bedingungen reagierten Neugeborene auf die Veränderung des Betonungsmusters. Erst dieser Befund lässt den sicheren Schluss zu, dass die Kinder in der Lage sind, Gemeinsamkeiten im Betonungsmuster aus einer Reihe segmental verschiedener Silbenketten zu extrahieren und dass sie beim Wechsel des Betonungsmusters tatsächlich auf die Betonungsmerkmale reagieren und nicht auf etwaige minimale segmentale Unterschiede, die mit dem Betonungsunterschied eines Lautsegments assoziiert sind.

Dies stellt eine wichtige Voraussetzung zur Ausnutzung rhythmischer Information im normalen sprachlichen Input dar, da dieser im allgemeinen auch diverse segmentale Strukturen enthält.

4.1.2.3 Grenzsignale

Christophe, Dupoux, Bertoncini und Mehler (1994) gingen der Frage nach, ob bereits Neugeborene sensitiv für prosodische Grenzsignale sind. In einem Experiment mit der *High-Amplitude Sucking* Methode stellten sie fest, dass 3 bis 4 Tage alte Säuglinge Silbensequenzen, die aus einem Wort herausgeschnitten worden waren (*mati* aus *mathématicien*), von Silbensequenzen diskriminieren, die aus zwei aufeinanderfolgenden Wörtern geschnitten worden waren (*mati* aus *pyjama tissé*). Eine phonetische Analyse der Silbenfolgen zeigte, dass die erste Silbe der Zweiersequenz eine längere Dauer und eine höhere akustische Energie aufwies, wenn ihre ursprüngliche Position wortfinal war gegenüber wortinternen Positionen. Offensichtlich sind schon Neugeborene in der Lage diese Dauer- und Energieunterscheide wahrzunehmen.

Die Wahrnehmung der prosodischen Grenzsignale für Intonationsphrasen und damit für Satz- und Teilsatzgrenzen, untersuchten Hirsh-Pasek, Kemler Nelson, Juszyck, Cassidy, Druss and Kennedy (1987). Sie boten Säuglingen Sprachpassagen, in die nach dem Einsprechen Pausen von einer 1 Sekunde geschnitten worden waren. Diese Pausen waren entweder an Satzgrenzen eingefügt worden oder an verschiedenen Positionen innerhalb von Konstituenten, (z. B. zwischen Kopula und Adjektiv). Die in die Studie einbezogenen Kinder hatten ein Alter zwischen 7 und 10 Monaten. In der Untersuchung mit dem *Headturn Preference* Paradigma zeigten sich längere Orientierungszeiten für die Textpassagen mit den Pausen an den Satzgrenzen. Die Autoren vermuten, dass diese Präferenz darin begründet ist, dass im ursprünglich eingesprochenen Sprachmaterial bereits prosodische Grenzsignale wie kurze Pausen, Veränderungen in der Grundfrequenz etc. in den Satzgrenzen enthalten waren. Die eingeschnittene Pause korrelierte damit mit den übrigen Grenzsignalen. In den Fällen, in denen die Pause satzintern eingeschnitten wurde, fehlten diese übrigen Grenzsignale natürlich. Danach scheinen Kinder schon sehr früh Wissen über die Korrelation

verschiedener phonetischer Merkmale in ihrer Funktion als Grenzsinal zu haben. Eine noch ausgeprägtere Präferenz für Textpassagen mit natürlichen Pausenpositionen an Satzgrenzen gegenüber Textpassagen, in die Pausen innerhalb von Sätzen eingeschnitten worden waren, zeigte sich bei der Verwendung von kindgerichteter Sprache als Stimulusmaterial (Kemler Nelson, Hirsh-Pasek, Jusczyk & Cassidy, 1989). Dieser Befund stützt erneut die vielfach gemachte Beobachtung einer besonders deutlichen prosodischen Markierung von Satzgrenzen als wesentlichem Merkmal von *Motherese* (als Überblick s. Fernald & McRoberts, 1996).

Jusczyk, Hirsh-Pasek, Kemler Nelson, Kennedy, Woodward und Piwoz (1992) konnten zeigen, dass etwas ältere Kinder auch bereits auf die prosodischen Grenzmarkierungen für kleinere Einheiten, nämlich Phrasen, beispielsweise die Grenze zwischen Nominal- und Verbalphrase, reagieren. Ihre Vorgehensweise war vollkommen analog zu der von Hirsh-Pasek und Mitarbeitern: sie präsentierten den Kindern Sprachpassagen, die aus einer natürlichen Interaktion einer Mutter mit ihrem zweijährigen Kind stammte. In diese Sprachpassagen wurden für die Präsentation im Experiment Pausen von einer Sekunde eingefügt, und zwar zwischen dem Subjekt und dem Verb eines Satzes oder nach dem Vollverb. Nur in der erstgenannten Position tritt die Pause an einer Stelle auf, an der prosodische Grenzsinal mit der Grenze einer syntaktischen Hauptkonstituenten konform gehen würden. Es zeigte sich, dass 9 Monate alte Kinder die Textpassagen mit der Pause vor dem Verb präferierten. Dieser Effekt zeigte sich bei 6 Monate alten Kindern dagegen noch nicht. Der Effekt war bei 9 Monate alten Kindern noch deutlicher, wenn statt der Passagen spontaner Sprache Passagen mit sehr langen Subjekten, die von einer Mutter vorgelesen worden waren, als Untersuchungsmaterial verwendet wurden. Die Effekte blieben stabil, unabhängig davon, ob die Mutter angewiesen worden war, die Texte in kind- oder erwachsenengerichteter Sprechweise zu artikulieren, wenn auch die Unterschiede in den Präferenzen bei der kindgerichteten Sprache größer waren als bei erwachsenengerichteter Sprache. Sie blieben auch bei der Verwendung *low-pass* gefilterten Materials erhalten.

Eine phonetische Analyse der verwendeten vorgelesenen Stimuluspassagen zeigte, dass in den Sätzen, in denen die Pause zwischen Subjekt und Verb eingefügt worden war, dieser Pause auch durchweg ein Abfall der Grundfrequenz vorausging, während

dies für die Pausen, die hinter das Verb geschnitten worden waren, nicht der Fall war. Zudem war die letzte Silbe vor der Pause zwischen Subjekt und Verb tendenziell länger als die letzte Silbe vor der Pause nach dem Vollverb. Beide Merkmale, nämlich der Grundfrequenzabfall sowie die Silbenlängung im Subjekt deuten darauf hin, dass hier beim Einsprechen des Materials vom Sprecher eine prosodische Grenze eingefügt worden war, so dass durch das Einfügen der Pause an dieser Stelle die prosodische Struktur erhalten blieb. In den übrigen Fällen fehlten diese Merkmale der prosodischen Grenzmarkierungen, so dass die Pause hier der ursprünglichen prosodischen Gliederung nicht entsprach. Da sich das gleiche Präferenzmuster auch zeigte, wenn den Kindern die Passagen *low-pass* gefiltert präsentiert wurden, scheint vor allem die Sensitivität der Kinder für die Zusammengehörigkeit der akustischen Merkmale der prosodischen Strukturierung (Grundfrequenzverlauf, Pausen, Dauerverhältnisse) ausschlaggebend für die Befunde zu sein. Möglicherweise hilft den Sprachlernern das frühe Wissen über die Zusammengehörigkeit dieser prosodischen Merkmale, in ihrem Input auftretende Pausen, die tatsächlich als Grenzsinal fungieren, von solchen zu unterscheiden, die durch Störungen des Sprachproduktionsprozesses (z.B. Wortfindungsschwierigkeiten) bedingt sind und damit nicht unbedingt an strukturell sinnvollen Positionen auftauchen.

Die Annahme prosodie-interner Gründe – und nicht wie vielfach angenommen der durch die „falsche“ Pausensetzung bedingte Widerspruch zwischen syntaktischer und prosodischer Struktur – als Quelle für dieses Präferenzmuster wird auch gestützt durch die Ergebnisse einer Studie von Gerken, Jusczyk und Mandel (1994). Sie variierten in ihrem Satzmaterial die prosodische Struktur bei gleichbleibender syntaktischer Struktur, in dem sie Sätze mit vollen Nominalphrasen als Subjekt und Sätze mit pronominalem Subjekt verwendeten. In ihrer syntaktischen Struktur sind beide Satztypen identisch: sowohl die volle Nominalphrase als auch das Pronomen bilden eine syntaktische Phrase. Sie unterscheiden sich jedoch in ihrer prosodischen Struktur: nur die volle Nominalphrase bildet eine eigenständige phonologische Phrase, das Pronomen als unbetontes Element kann demgegenüber keine selbständige phonologische Phrase bilden, sondern bildet zusammen mit dem folgenden Verb eine einzige phonologische Phrase. Wird in diesen beiden Satztypen eine Pause zwischen Subjekt und Verb eingefügt, d. h. an syntaktisch identische Positionen, lässt diese Pause die phonologische Phrasierung der Sätze mit voller Subjekt-Nominalphrase intakt, stört

jedoch die phonologische Phrasierung der Sätze mit pronominalem Subjekt. Die Ergebnisse der Untersuchung von Gerken und Mitarbeitern mit 9 Monate alten Kindern reflektiert genau diesen phonologischen Unterschied zwischen den Satztypen: bei den Sätzen mit pronominalem Subjekt zeigte sich kein Unterschied zwischen den Reaktionen auf die Sätze mit einer Pause zwischen Subjekt und Verb und den Reaktionen auf die Sätze mit phraseninterner Pause. Demgegenüber präferierten die Kinder bei den Sätzen mit voller Subjekt-NP, diejenigen mit Pause zwischen Subjekt und Verb gegenüber denjenigen mit phraseninterner Pause. Diese Ergebnisse belegen deutlich, dass die Befunde der drei zuletzt vorgestellten Studien noch kein syntaktisches Wissen der Kinder belegen, sondern sie zeigen eine sehr früh vorhandene Sensitivität gegenüber prosodischen Grenzmarkierungen in ihrem sprachlichen Input. Damit steht den Kindern ein erstes Mittel zu Verfügung, den zu verarbeitenden Sprachstrom zu strukturieren. Da prosodische Einheiten vielfach mit syntaktischen Einheiten korrelieren, kann die prosodische Strukturierung auch einen ersten Einstieg in die syntaktische Strukturierung des Inputs liefern. Eine prosodische Segmentierung würde dem Kind Einheiten liefern, die in den meisten Fällen auch syntaktische Einheiten bilden, und somit für die Extraktion syntaktischer Regularitäten aus dem Input relevante Informationseinheiten darstellen.

4.2 Die Ausnutzung prosodischer Merkmale für die Gliederung des Signals

4.2.1 Verarbeitung prosodischer Information und Erwerb lexikalischer Information

4.2.1.1 Wortsegmentierung

Anders als für die Sprachverarbeitung beim Erwachsenen muss für den frühen Spracherwerb angenommen werden, dass lexikalische Information für die Lösung des Segmentierungsproblems zunächst keine Rolle spielen kann, da die Segmentierung dem

Aufbau zumindest von rudimentärem lexikalischem Wissen vorausgehen muß. Dies legt die Vermutung nahe, daß es zunächst die rhythmische Information ist, die die Kinder für die Segmentierung des Signals ausnutzen.

Vorliegende empirische Befunde weisen übereinstimmend darauf hin, dass Säuglinge bereits in der zweiten Hälfte ihres ersten Lebensjahres rhythmische Merkmale zur Gliederung akustischer Signale nutzen. So fanden Morgan und Saffran (Morgan, 1994, 1996; Morgan & Saffran, 1995), dass Kinder ab dem Alter von 6 Monaten langsamer auf ein Störgeräusch reagieren, wenn dieses Störgeräusch innerhalb einer rhythmisch strukturierten Einheit auftritt, als wenn das Störgeräusch außerhalb der rhythmischen Einheit auftritt. Dieses Ergebnis deutet darauf hin, dass eine Silbenfolge, die eine rhythmische Struktur aufweist, eine Wahrnehmungseinheit darstellt.

Eine deutliche Tendenz, dass 9 Monate alte englische Kinder trochäische Sequenzen aus längeren Silbenketten extrahieren, fanden Echols, Crowhurst und Childers (1997). In der Familiarisierungsphase ihres Experiments präsentierten sie prosodisch strukturierte Ketten von Pseudowörtern, die sie aus englischen Phrasen durch Ersetzung aller Morpheme durch Unsinnssilben erzeugt hatten. Diese Stimuli enthielten trochäische und auch jambische Pseudowörter. In der Testphase wurden den Kindern diese Pseudowörter neben anderen in der Familiarisierungsphase nicht vorgekommenen präsentiert. Es zeigte sich, dass die Kinder in der Testphase die trochäischen Pseudowörter wiedererkannten, nicht aber die jambischen.

Der Frage, ob rhythmische Information bei der Segmentierung längerer natürlichsprachlicher Äußerungen eine Rolle spielt, gingen Jusczyk und Mitarbeiter in einer Untersuchungsreihe nach (Jusczyk, Houston & Newsome, 1999). Nach der Hypothese der metrischen Segmentierung, bei der im Englischen eine starke Silbe als Wortanfang und die folgenden schwachen Silben als Bestandteil des Wortes behandelt würden, wären für Wörter, die nicht mit einer starken Silbe beginnen, Fehlsegmentationen zu erwarten. Dies wurde unter Verwendung des *Headturn Preference* Paradigmas überprüft. Dabei wurden den Kindern in der Familiarisierungsphase zunächst isolierte Wörter präsentiert. In der darauf folgenden Testphase wurden den Kindern kurze Texte dargeboten, von denen ein Teil die vorher familiarisierten Wörter enthielt und der andere Teil keins der familiarisierten Wörter enthielt. Wenn es sich bei den familiarisierten Wörtern um zweisilbige Wörter im

trochäischen Muster handelte, zeigten Kinder im Alter zwischen 7 und 8 Monaten längere Orientierungszeiten für die Texte, die die kritischen Wörter enthielten, nicht aber wenn es sich bei den familiarisierten Wörtern um zweisilbige Wörter im jambischen Muster handelte. Auch bei umgekehrtem Vorgehen, d. h. wenn den Kindern in der Familiarisierungsphase die Texte und in der Testphase die isolierten Wörter präsentiert wurden, zeigte sich ein Wiedererkennungseffekt nur für die trochäischen nicht aber für die jambischen Wörter.

Die Vermutung, dass trochäische Wörter in diesem Alter bereits als Einheiten repräsentiert werden, jambische dagegen nicht, wird durch weitere Untersuchungsbefunde aus der selben Studie belegt. Eine Familiarisierung mit der starken Silbe eines trochäischen Wortes (z.B. *ham* aus *hamlet* entweder im Text oder isoliert) führte nicht zu einer Erkennungsreaktion auf das gesamte trochäische Wort und eine Familiarisierung mit einem trochäischen Wort führte nicht zu einer Erkennungsreaktion bei Präsentation der isolierten starken Silbe. Anders ist wiederum das Bild für die jambischen Wörter: eine Familiarisierung mit der isolierten Silbe eines jambisches Wortes (*tar* aus *guitar*) führt zur Reaktion auf das jambische Wort im Kontext und eine Familiarisierung mit dem jambischen Wort führt zur Reaktion auf die isolierte Silbe. Nach diesen Befunden repräsentieren englische Kinder ab einem Alter von ca. 7 Monaten trochäische Wörter als Einheit, wobei die starke Silbe nicht ohne weiteres aus dieser Einheit herausgelöst wird. Demgegenüber scheint der Ansatz der betonten Silbe im Jambus ein potentieller Segmentierungspunkt zu sein.

Nach diesem Prinzip ist die Segmentation des sprachlichen Inputs in Einheiten zu erwarten, die nicht unbedingt den lexikalischen Einheiten der Standardsprache entsprechen. Genau dies wird durch ein weiteres Experiment der Reihe belegt. Hier wurden in der Familiarisierungsphase Texte mit jambischen Wörtern präsentiert, denen stets die selbe schwache Silbe folgte (z.B. *the guitar is*). In der anschließenden Testphase zeigten die Kinder einen Wiedererkennungseffekt für ein aus der starken und der folgenden schwachen Silbe gebildetes Pseudowort (z. B. *taris*). Dies deutet darauf hin, dass es zu einer Fehlsegmentation der Textpassagen kam und – wahrscheinlich gestützt durch die Tatsache, dass diese Segmente mehrfach im Text vorkamen - als Konsequenz eine nicht standardsprachliche Repräsentationseinheit gebildet wurde. Jusczyk und Mitarbeiter führten einen Teil der oben geschilderten Experimente auch

mit älteren Kindern im Alter zwischen 10 und 11 Monaten durch. In dieser Altersgruppe fanden sich keine Unterschiede mehr zwischen der Verarbeitung von jambischen und trochäischen Wörtern: die Kinder erkannten auch jambische Wörter erfolgreich wieder und reagierten nicht mehr auf einzelne Silben eines jambischen Wortes.

Mit einer anderen Methode zeigten auch Myers und Mitarbeiter, dass Kinder im Alter von ungefähr 11 Monaten sowohl trochäische als auch jambische Wörter im Kontext als Einheit erkennen (Myers, Jusczyk, Kemler Nelson, Charles Luce, Woodward & Hirsh-Pasek, 1996). Die Kinder dieser Studie zeigten eine Präferenz für Texte, in die Pausen zwischen Wörtern eingefügt worden waren, gegenüber Texten, in die Pausen zwischen den Silben eines Wortes eingefügt worden waren. Dieses Muster zeigte sich unabhängig davon, ob die betroffenen Wörter ein trochäisches oder ein jambisches Betonungsmuster aufwiesen. In diesem Alter scheinen rhythmische Merkmale also nicht mehr allein ausschlaggebend für die Segmentierung gesprochener Sprache zu sein.

Zusammenfassend lässt sich also sagen, dass englische Säuglinge im Alter zwischen 7 und 9 Monaten eine Segmentierungsstrategie anwenden, die deutlich vom rhythmischen Muster der Umgebungssprache geprägt ist. Es soll nun der Frage nachgegangen werden, ob sich auch bei deutschen Kindern eine solche deutliche Tendenz zur trochäischen Segmentierung nachweisen lässt. Das Englische und das Deutsche sind sich in Bezug auf rhythmische Merkmale sehr ähnlich: beide Sprachen werden zu den akzentzählenden Sprachen gezählt und der Trochäus gilt auch im Deutschen als das dominante Betonungsmuster (Wiese, 1996).

Die Vermutung einer Trochäusdominanz im Deutschen wurde zunächst anhand einer Analyse aller zweisilbigen Inhaltswörter (Nomen, Verben, Adjektive) aus dem deutschsprachigen Korpus der Celex Datenbank (Baayen, Piepenbrock & van Rijn, 1993) überprüft. Von den ca. 10.000 Worttypen dieser Kategorien wiesen 85% eine starke erste Silbe auf. Berücksichtigt man zusätzlich die Wortfrequenz, so ist davon auszugehen, dass ca. 91% der zweisilbigen Inhaltsworttoken das trochäische Betonungsmuster aufweisen. Diese Zahl ist beinahe identisch zu den von Cutler und Carter (1987) gefundenen Verhältnissen in einem englischen Korpus gesprochener Sprache, in dem 91% aller Inhaltswörter eine initiale starke Silbe aufwiesen. Diese

Daten sprechen dafür, dass Englisch und Deutsch in Bezug auf die Trochäusdominanz tatsächlich vergleichbar sind.

4.2.1.2 Experiment 1: Detektion trochäischer Zweisilbler bei deutschen Säuglingen

Falls die rhythmusbasierte Segmentationsstrategie ein universales Verarbeitungsprinzip darstellt, sollten sich bei deutschen Kindern aufgrund der rhythmischen Ähnlichkeit zwischen Englisch und Deutsch die gleichen Effekte zeigen, die in den obigen Studien für englische Kinder gefunden wurden. Diese Hypothese wurde in einer Untersuchung zu frühen Wortsegmentierungsleistungen überprüft, d. h. es wurde getestet, ob deutsche und englische Kinder in Bezug auf ihre Wortsegmentierungsleistungen bei trochäischen Wörtern vergleichbar sind.

Probanden

In die Untersuchung einbezogen wurden 28 Kinder im Alter zwischen 8 und 9 Monaten. Das Durchschnittsalter betrug 8 Monate 26 Tage bei einer Streubreite von 8 Monate 1 Tag bis 9 Monate 12 Tage. In der Stichprobe befanden sich 12 Mädchen und 16 Jungen. Alle Kinder waren termgeboren und stammten aus monolingualen, deutschsprachigen Elternhäusern. Zusätzlich wurden 13 Kinder getestet, deren Ergebnisse wegen Experimentabbrüchen oder zu kurzen mittleren Orientierungszeiten nicht in die Analyse eingingen.

Material

Als Stimuluswörter dienten vier trochäische Wörter (Balken, Pinsel, Kurbel, Felsen). Die Wörter waren nach folgenden Kriterien ausgewählt:

1. Die starke Silbe entspricht einer im Deutschen existierenden Wortform.
2. Die schwache Silbe der trochäischen Wörter enthält einen auslautenden Konsonanten.
3. Die Wörter unterscheiden sich möglichst weitgehend in ihrem segmentalen Inventar.

Für jedes der Wörter wurde eine Textpassage bestehend aus sechs einfachen Sätzen konstruiert. In jedem der Sätze einer Passage kam das kritische Wort genau einmal vor, wobei darauf geachtet wurde, dass die Positionen des kritischen Wortes in den verschiedenen Sätzen unterschiedlich waren. Der folgende Text ist ein Beispiel für eine der verwendeten Passagen:

*Der **Balken** lag quer über dem tiefen Abgrund. Es war ein sehr breiter und stabiler **Balken**. Über diese **Balken** spazierten die Wanderer auf die andere Seite des Tales. Manchmal liefen auch Kinder den **Balken** entlang. Auf dem dicken **Balken** zu spielen war sehr gefährlich. Zum Glück ist noch nie ein Kind vom **Balken** gefallen.*

Das gesamte Untersuchungsmaterial wurde von einer Sprecherin des Deutschen eingesprochen, die angewiesen wurde, das Material in einer lebhaften, prosodisch variablen Sprechweise zu artikulieren, so als ob sie zu einem Kind sprechen würde. Die einzelnen Wörter wurden von der Sprecherin ca. 40 mal mit leicht variierender Intonation eingesprochen. Das gesamte Material wurde anschließend digitalisiert.

Aus den digitalisierten Sprachdateien wurde für jedes Wort eine Sequenz von 30 Tokens des jeweiligen Wortes angefertigt, wobei zwischen den einzelnen Token eine Pause von 600 ms geschnitten wurde. Die digitalisierten Textpassagen wiesen Dauerwerte um die 19 Sekunden auf (*Balken*-Text: 18,65 s; *Pinsel*-Text: 19,19 s; *Felsen*-Text: 19,34 s; *Kurbel*-Text: 19,03 s).

Durchführung

Verwendet wurde das *Headturn Preference* Paradigma mit Familiarisierungsphase wie in Abschnitt 2.3.1 beschrieben. In der Familiarisierungsphase wurden den Kindern die Sequenzen für zwei der vier Wörter alternierend präsentiert, bis sie jedes Wort für mindestens 30 s gehört hatten. Anschließend wurden ihnen die Textpassagen für alle vier Wörter eines rhythmischen Musters je viermal in durchmischter Reihenfolge präsentiert.

Die Orientierungszeiten wurden von Beginn der sprachlichen Präsentation an gemessen. Drehte das Kind den Kopf für mehr als zwei Sekunden vom jeweiligen

Lautsprecher weg, wurde der Durchgang abgebrochen und der nächste mit dem Blinken der grünen zentralen Lampe gestartet. Drehte das Kind den Kopf von der Präsentationsseite weg, schaute jedoch innerhalb von zwei Sekunden zurück, wurde der Durchgang fortgesetzt, die Zeit des Wegschauens ging jedoch nicht in die gemessene Orientierungszeit ein.

Ergebnisse und Diskussion der Ergebnisse

Für jedes Kind wurden zunächst die absoluten Orientierungszeiten für jeden der vier Texte ermittelt. Dann wurden die Orientierungszeiten jeweils für die Texte mit familiarisierten Wörtern und die Texte ohne familiarisierte Wörter zusammengefasst und gemittelt.

Für die Texte, die eins der familiarisierten Wörter enthielten, zeigten sich mit durchschnittlich 6,17 s signifikant längere Orientierungszeiten als für die Texte, die keins der familiarisierten Wörter enthielten (5,31 s). Dieser Unterschied in den Orientierungszeiten ist statistisch signifikant ($t(df = 27) = 3.08; p < 0.01$).

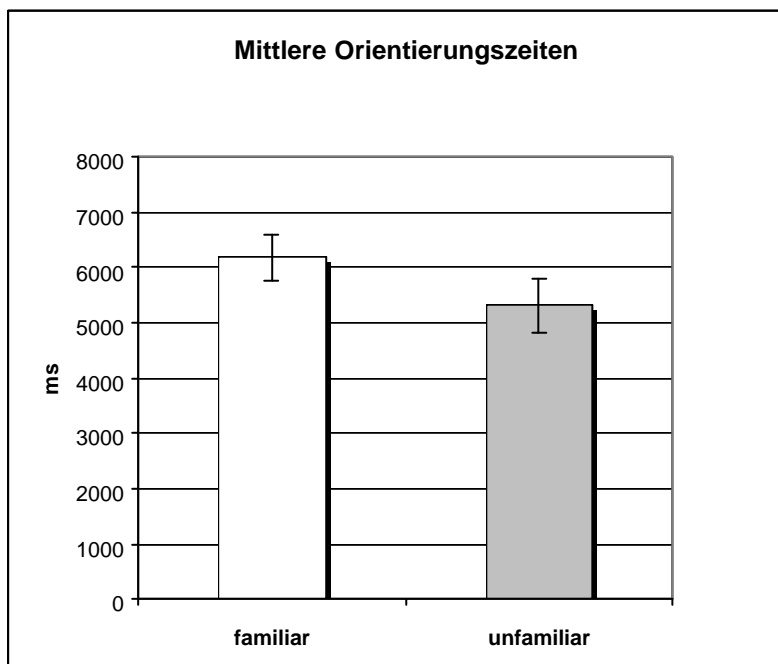


Abbildung 1: Mittlere Orientierungszeiten für Textpassagen mit familiarisierten und ohne familiarisierte Wörter aus Experiment 1

Dieses Ergebnis zeigt, dass auch deutschsprachigen Kinder im Alter von ca. 8 Monaten in der Lage sind, zweisilbige trochäische Wörter, die sie zuvor isoliert gehört haben, in einer kurzen Textpassage zu entdecken. Dies deutet daraufhin, dass die deutschen Kinder, wie ihre englischsprachigen Altersgenossen, anfänglich eine metrische Segmentierungsstrategie verwenden.

Wenn die Ergebnisse für die deutschsprachigen Kinder tatsächlich die Anwendung einer rhythmischen Segmentierungsstrategie reflektieren und die deutschsprachigen Kinder auf den gleichen Typ von Information reagieren, wie ihre englischsprachigen Altersgenossen, dann sollten sich englische Kinder gegenüber dem deutschsprachigen Stimulusmaterial in genau der gleichen Weise verhalten, wie für die deutschen Kinder gezeigt. Dass Kinder die innerhalb ihrer eigenen Sprache entwickelte Segmentationsstrategie auch auf andere Sprachen übertragen – und zwar auch in einem Alter, in dem sie Unterschiede zwischen ihrer Muttersprache und einer Fremdsprache bereits an segmentalen Unterschieden erkennen, zeigen Befunde von Houston, Jusczyk, Kuijpers, Coolen und Cutler (2000). Sie konnten zeigen, dass englische 9 Monate alte Kinder wie ihre niederländischen Altersgenossen in der Lage sind, niederländische trochäische Wörter in niederländischen Textpassagen wiederzuerkennen. Aufgrund der rhythmischen Ähnlichkeit zwischen dem Niederländischen und dem Englischen führt hier die Übertragung der muttersprachlich erfolgreichen Segmentierungsstrategie in eine andere Sprache ebenfalls zu einer erfolgreichen Erkennung der Wörter. Frühere Untersuchungen gerade für diese beiden Sprachen haben gezeigt, dass 9 Monate alte Kinder sie anhand segmentaler Unterschiede differenzieren können (Jusczyk, Friederici, Wessels, Svenkerud & Jusczyk, 1993).

Nach diesen Befunden ist also zu erwarten, dass englische Kinder auch in der Lage sind, deutsche trochäische Wörter in deutschen Textpassagen zu detektieren. Diese Hypothese wurde in einer Untersuchung an englischen Kindern im Labor von Peter Jusczyk, Johns Hopkins University Baltimore überprüft.

4.2.1.3 Experiment 2: Detektion trochäischer Zweisilbler in deutschen Stimuli durch englische Säuglinge

Probanden

An der Untersuchung nahmen 30 Säuglinge (14 Mädchen, 16 Jungen) teil, die aus monolingualen englischsprachigen Elternhäusern aus Baltimore und Umgebung stammten. Die Kinder waren ebenfalls zwischen 8 und 9 Monate alt und wiesen ein durchschnittliches Alter von 9 Monaten und 4 Tagen auf (Spannbreite: 8 Monate 20 Tage bis 9 Monate 13 Tage).

Material und Durchführung

Das verwendete Material ist identisch mit dem im deutschen Experiment verwendeten. Die Durchführung des *Headturn Preference* Paradigmas in den Laboren in Potsdam und Baltimore ist weitestgehend gleich. Die experimentelle Vorgehensweise während der Familiarisierungs- und während der Testphase war die gleiche wie die mit den deutschen Kindern.

Ergebnisse und Diskussion der Ergebnisse

Wiederum wurden die Orientierungszeiten jedes Kindes für jeden Text ermittelt und daraus die mittleren Orientierungszeiten für die Texte mit familiarisierten und die Text ohne familiarisierte Wörter berechnet. Die englischen Säuglinge zeigten eine durchschnittliche Orientierungszeit von 7,13 s für die Texte mit familiarisierten Wörtern und von 6,19 s für die Texte ohne familiarisierte Wörter. Dieser Unterschied ist statistisch signifikant ($t(df=29) = 2.448; p < 0.05$).

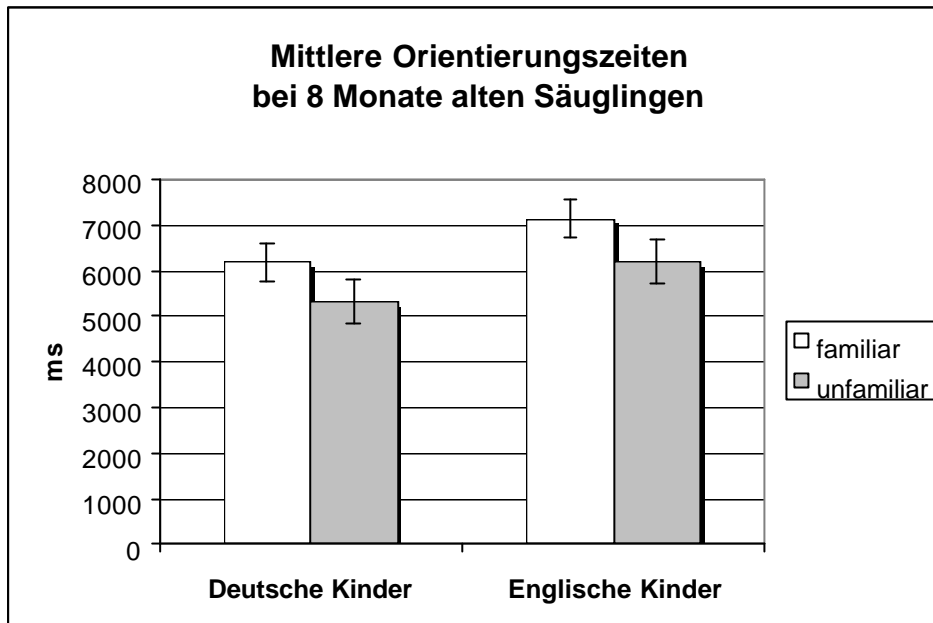


Abbildung 2: Mittlere Orientierungszeiten für Textpassagen mit familiarisierten und ohne familiarisierte Wörter aus den Experimenten 1 und 2

Nach diesen Ergebnissen verhalten sich also englischsprachige Kinder bei der Präsentation deutscher trochäischer Wörter in deutschsprachigen Texten genauso wie deutschsprachige Kinder und auch genauso, als ob ihnen Stimulusmaterial ihrer eigenen Sprache präsentiert wird. Unabhängig von speziellen segmentalen Charakteristika der sprachlichen Reize sind sie in der Lage rhythmische Einheiten, speziell trochäische Einheiten aus dem Sprachfluss herauszusegmentieren. Dies zeigt, dass bis zu einem gewissen Grad eine rein prosodische Segmentierung tatsächlich zu einer Extraktion wortentsprechender Einheiten aus dem Input führen kann, ohne dass hierzu Wissen über segmentale oder lexikalische Besonderheiten des sprachlichen Materials erforderlich ist.

Zusammengenommen zeigen die bisher vorliegenden Ergebnisse mit englischen, niederländischen und deutschsprachigen Kindern erstaunliche Parallelen zu den Befunden mit Erwachsenen, wo sich übereinstimmend zeigte, dass die Segmentierungsstrategie der Muttersprache auch bei der Verarbeitung nicht-muttersprachlicher Reize angewendet wird (Cutler et al., 1986). Offensichtlich verhalten sich bereits 8 Monate alte Säuglinge in dieser Hinsicht wie erwachsene Hörer. Es wäre interessant zu sehen, ob diese Segmentierungsstrategie auch von Kindern schon auf

Sprachen eines anderen Rhythmustypen übertragen wird. Gezielte Untersuchungen hierzu fehlen jedoch bislang.

Zwar stützen die sprachvergleichenden Befunde die innerhalb der rhythmischen Segmentierungshypothese gemachte Annahme, dass sich bereits im frühen Spracherwerb der rhythmische Segmentierungsmechanismus an die in der Muttersprache vorgefundenen phonologischen Verhältnisse adaptiert: in den untersuchten akzentzählenden Sprachen zeigt sich übereinstimmend die frühe Fähigkeit, trochäische Wörter aus dem Sprachfluss herauszusegmentieren. Es bleibt jedoch noch die Frage offen, ob diese Strategie bereits eine Adaptation an die in der Muttersprache vorgefundenen rhythmischen Gegebenheiten reflektiert, oder ob sie eine grundsätzliche Vorliebe von Säuglingen für trochäische Muster widerspiegelt. Theoretisch scheint die erste Alternative plausibler, da – falls die zweite korrekt wäre – der Erwerb einer Sprache, in der der Trochäus nicht das dominante rhythmische Muster darstellt, grundsätzlich schwieriger sein sollte als der einer trochäus-dominanten Sprache. Anhand der bisherigen Befunde lässt sich zwischen diesen Alternativen allerdings nicht eindeutig entscheiden. Dieser Frage wird im nächsten Abschnitt weiter nachgegangen.

4.2.2 Die Trochäuspräferenz: Resultat einer Anpassung an die Umgebungssprache?

Bislang ist die Datenlage zur Frage, ob die für Säuglinge im Alter von ca. 8 bis 9 Monaten beobachtete Trochäuspräferenz eine Anpassung an die prosodischen Verhältnisse der Zielsprache darstellt oder ob sie eine universelle Ausgangspräferenz darstellt, vergleichsweise dünn. Erste Hinweise, die für die Entwicklung einer Trochäuspräferenz sprechen, finden sich in einer Arbeit von Jusczyk, Cutler und Redanz (1993). Sie präsentierten englischen Kindern Listen von Wörtern, wobei jede Liste entweder nur trochäische oder nur jambische Wörter des Englischen enthielt. Bei 9 Monate alten Säuglingen zeigte sich eine signifikante Präferenz für die Listen trochäischer Wörter, während sich bei 6 Monate alten Kindern kein Unterschied in den Orientierungszeiten für trochäische und jambische Wörter fand. Die Tatsache, dass die 9 Monate alten Kinder die gleiche Präferenz zeigten, wenn ihnen das Untersuchungsmaterial *low-pass* gefiltert präsentiert wurde, spricht dafür, dass

tatsächlich die prosodischen Unterschiede zwischen den jambischen und den trochäischen Wörtern für diese Präferenz verantwortlich war und nicht beispielsweise Unterschiede im segmentalen Aufbau. Dieses Präferenzmuster wurde in einer späteren Arbeit von Mattys, Jusczyk, Luce und Morgan (1999), in der statt realer Wörter zweisilbige Pseudowörter verwendet wurden, bestätigt. Auch hier zeigten 9 Monate alte englische Kinder eine Präferenz für trochäische Zweisilbler gegenüber jambischen Zweisilblern.

Auch in der schon erwähnten Studie von Echols und Mitarbeitern (1997) wurden in einem der Experimente die Reaktionen auf rhythmische Information bei sechssmonatigen und bei neunmonatigen Säuglingen verglichen. In dieser Untersuchung wurden den Probanden in der Familiarisierungsphase dreisilbige Sequenzen, deren mittlere Silbe betont war, präsentiert. Während der Testphase wurde in diese Sequenzen eine Pause integriert, die entweder vor der betonten Silbe (be_GUdi) oder nach der betonten Silbe (beGU_di) auftrat. Wenn die Kinder eine trochäische Präferenz für die Strukturierung dieser Reize haben, sollten sie unterschiedlich auf diese Reize reagieren, da der Einschub der Pause vor der betonten Silbe den Trochäus als rhythmische Einheit erhält, während der Einschub der Pause nach der betonten Silbe die trochäische Einheit unterbricht. Tatsächlich zeigten 9 Monate alte Kinder eine Präferenz für die Sequenzen mit Pause vor der betonten Silbe, während die 6 Monate alten Kinder keinen Unterschied in den Orientungszeiten für die verschiedenen Pausenpositionen zeigten.

Auch Morgan (1996) ging der Frage nach etwaigen Unterschieden in der Trochäuspräferenz bei 6 und 9 Monate alten Kinder nach. Ausgangspunkt seiner Untersuchungen war der Befund von Morgan und Saffran (1995), nach dem neunmonatige Säuglinge trochäische und jambische Zweisilbler in gleicher Weise als kohärente Einheiten repräsentieren. Eine mögliche Erklärung dieses Widerspruchs zu den anderen Ergebnissen, nach denen 9-Monate alte englische Säuglinge eine Präferenz für trochäische Einheiten zeigen, könnte in methodischen Unterschieden zwischen den Studien liegen. In der von Morgan und Saffran verwendeten *Noise Detection* Technik wurden den Säuglingen während der Konditionierungs- und Testphase des gesamten Experiments immer dieselben Zweisilbler geboten. Die einzige Variation bestand in zwei verschiedenen Füllsilben, die vor oder nach dem rhythmisch strukturierten

Zweisilbler geboten wurden. Da der kritische Reiz, auf den die Kinder reagieren mussten, ein Störgeräusch war, das zwischen oder innerhalb dieser Zweisilbler auftrat, wurden die Silbensequenzen als Hintergrundstimulus während des gesamten Experiments, d. h. auch in der Trainingsphase verwendet. Bevor die eigentliche Testphase des Experiments begann, hatten die Kinder die kritischen Zweisilbler also schon sehr häufig gehört. Diese häufige Präsentation könnte dazu geführt haben, dass sich eventuell am Beginn des Experiments bestandene Unterschiede zwischen den trochäischen und den jambischen Silbensequenzen nivelliert haben, d. h. die Kinder haben im Laufe des Experiments gelernt, dass auch die jambischen Zweisilbler eine Einheit darstellen. Hier könnte auch die Tatsache eine Rolle spielen, dass die Silben in immer der gleichen Reihenfolge geboten wurden, ein Merkmal, das die Bildung von Einheiten ebenfalls stützt. Um dieser Vermutung eines Lerneffekts weiter nachzugehen, führte Morgan (1996) eine Reihe weiterer Untersuchungen durch. Im Unterschied zur ersten Untersuchungsreihe wurden hier in der Testphase neben Zweisilbern, die den Kindern zuvor schon mehrfach präsentiert worden waren, auch Zweisilber benutzt, die im Experiment vorher noch nicht verwendet worden waren. Es zeigte sich, dass 9 Monate alte Kinder beim trochäischen Muster keinen Unterschied zwischen den "alten" und den "neuen" Reizen machten: sie reagierten in beiden Fällen langsamer auf ein Störgeräusch, das innerhalb eines trochäischen Zweisilblers präsentiert wurde als auf ein Störgeräusch, das außerhalb des Zweisilblers präsentiert wurde. Im jambischen Muster zeigte sich dieser Effekt dagegen nur für die "alten" Reize, nicht aber für die "neuen". Bei den 6 Monate alten Kindern fand sich dieser Unterschied zwischen den beiden rhythmischen Mustern dagegen nicht: sie zeigten in beiden Fällen sowohl bei den "alten" als auch bei den "neuen" Reizen langsamere Reaktion auf sequenzinterne Störgeräusche als auf sequenzexterne Störgeräusche.

Genau das gleiche Muster ergab sich auch bei einem weiteren Experiment mit leicht veränderten Vorgehensweise: hier hörten die Kinder einen Zweisilbler in der Trainingsphase des Experiments in einem konstanten rhythmischen Muster, entweder als Trochäus oder als Jambus. In der Testphase des Experiments wechselte der Rhythmus der gleichen Silbenabfolge vom Trochäus zum Jambus bzw. vom Jambus zum Trochäus. Bei 9 Monate alten Kindern blieb die verlangsamte Reaktion auf das Störgeräusch beim Wechsel vom Jambus zum Trochäus erhalten, nicht aber beim

Wechsel vom Trochäus zum Jambus. Dagegen reagierten 6 Monate alte Kinder auf den Rhythmuswechsel in gleicher Weise: in beiden Fällen verschwand die Verlangsamung der Reaktion.

Übereinstimmend zeigt sich also in allen vorgestellten Studien bei 9 Monate alten Kindern eine Trochäuspräferenz bzw. eine Präferenz, sprachliche Stimuli in trochäische Füße zu segmentieren. Dagegen scheinen 6 Monate alte Säuglinge noch keine Präferenz für den einen oder anderen Rhythmustypen zu haben, sie setzen offensichtlich rhythmische Information jedweder Art für die Verarbeitung sprachlicher Information ein. Dies spricht für die Annahme, dass die im neunten Lebensmonat bei englischen Kindern gezeigte Trochäuspräferenz tatsächlich bereits eine Adaptation an die Trochäusdominanz der Zielsprache darstellt. Im Alter von 6 Monaten scheint diese einzelsprachspezifische Festlegung noch nicht stattgefunden zu haben. Allerdings lassen sich die vorliegenden Daten noch nicht mit Sicherheit als Evidenz dafür interpretieren, dass Kinder im Alter von 6 Monaten noch kein Wissen über das in ihrer Sprache vorherrschende Betonungsmuster haben. Zu einem fanden Echols et al. (1977) und Jusczyk et al. (1992) bei 6 Monate alten Kindern keinerlei signifikante Unterschiede in den Orientierungszeiten für trochäische und jambische Muster. Nulleffekte lassen jedoch nicht darauf schließen, dass die Kinder tatsächlich noch nicht über dieses Wissen verfügen. Die Resultate von Morgan (1996) lassen ebenso noch verschiedene Interpretationsmöglichkeiten offen: es wäre denkbar, dass bei 6 Monate alten Kindern die Trochäuspräferenz noch nicht so gefestigt ist wie bei 9 Monate alten und daher leichter durch eine konzentrierte Präsentation jambischer Muster überlagert werden kann.

Um der Frage nach etwaigen Veränderungen und dem Einfluss sprachspezifischer Eigenschaften auf rhythmische Präferenzen weiter nachzugehen, wurde eine Untersuchung mit deutschsprachigen Säuglingen im Alter zwischen 3 und 6 Monaten durchgeführt, um zu überprüfen, ob sich in diesem Altersbereich bereits Veränderungen in den Reaktionen auf trochäische und auf jambische Zweisilbler nachweisen lassen.

4.2.2.1 Experiment 3: Trochäuspräferenz bei 6 Monate alten Säuglingen

In dieser Untersuchung wurde zunächst an einer Gruppe von 6 Monate alten Kindern überprüft, ob sich bei der Verwendung segmental sehr einfacher Stimuli eine Präferenz für den trochäischen Fuß nachweisen lässt.

Probanden

An der Untersuchung nahmen insgesamt 28 Kinder im Alter von 6 Monaten aus monolingualen deutschsprachigen Elternhäusern teil. Alle Kinder waren termgeboren. Vier Kinder wurden aufgrund von Experimentabbrüchen aus der Auswertung ausgeschlossen, so dass die endgültige Stichprobe aus 24 Kindern bestand. Diese Kinder wiesen ein durchschnittliches Alter von 6 Monaten und 12 Tagen auf. Die Spannweite lag zwischen 6 Monaten und einem Tag und 6 Monaten und 29 Tagen. Die Gruppe umfasste 11 Mädchen und 13 Jungen.

Material

Als Untersuchungsmaterial wurde die zweisilbige Sequenz *gaba* verwendet. Diese wurde von einer Muttersprachlerin des Deutsch insgesamt 10 mal im trochäischen Betonungsmuster und 10 mal im jambischen Betonungsmuster gesprochen. Aus diesen je 10 Token wurden pro Rhythmusmuster 6 verschiedene Sequenzen hergestellt, die entweder nur trochäische oder nur jambische Zweisilbler enthielten. Jede Sequenz enthielt insgesamt 16 Token der trochäischen bzw. jambischen Silbensequenzen mit einer Pause von 600 ms zwischen den einzelnen Token. Die Sequenzen wiesen insgesamt eine Dauer von ca. 18 s auf.

Durchführung

Das Experiment wurde mit der *Headturn Preference* Methode durchgeführt. Den Kindern wurden die 12 Untersuchungssequenzen zwischen den beiden Rhythmustypen alternierend präsentiert. Zwischen den Kindern wurde dabei variiert, ob mit einer Trochäus- oder einer Jambusequenz begonnen wurde, und auf welcher Seite der erste Stimulus präsentiert wurde. Die ersten beiden Sequenzen wurden als Übungsdurchgänge verwendet, in denen die Kinder eine Assoziation zwischen ihrer Kopfdrehung und der Dauer der Lautpräsentation herstellen und sich an den

Untersuchungsablauf gewöhnen sollten. Diese Übungsdurchgänge wurden nicht in die Analyse der Ergebnisse einbezogen.

Ergebnisse und Diskussion der Ergebnisse

Für jedes Kind wurden die Orientierungszeiten für die verschiedenen trochäischen und die Orientierungszeiten für die verschiedenen jambischen Untersuchungsdurchgänge pro Rhythmustyp gemittelt. Für die trochäischen Sequenzen zeigte sich eine durchschnittliche Orientierungszeit von 7,97 s. Die durchschnittliche Orientierungszeit für die Jamben betrug 6,71 s. Nach einem t-Test für abhängige Stichproben war dieser Unterschied statistisch signifikant ($t(df=23) = 2,16; p < 0.05$).

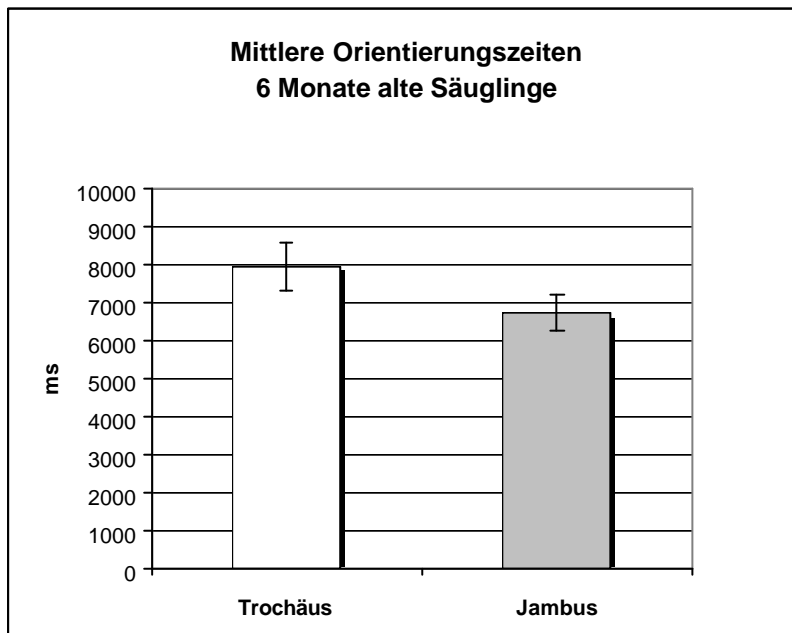


Abbildung 3: Mittlere Orientierungszeiten für trochäische und jambische Zweisilbler aus Experiment 3

Nach diesem Datenmuster zeigt sich also bereits bei 6 Monate alten deutschen Säuglingen eine deutliche Präferenz für das trochäische Betonungsmuster. Dieses Präferenzmuster lässt zwei Interpretationsmöglichkeiten offen: entweder es reflektiert eine generelle Vorliebe für den Trochäus, der möglicherweise sehr grundlegende Wahrnehmungsmechanismen reflektiert, oder diese Präferenz ist bereits ein Adaptation an das im Deutschen vorherrschende Betonungsmuster. Um einer Klärung zwischen

diesen beiden Interpretationen näher zu kommen, wurde die gleiche Untersuchung mit einer Gruppe von 3 Monate alten Säuglingen durchgeführt.

4.2.2.2. Experiment 4: Trochäuspräferenz bei 3 Monate alten Säuglingen

Probanden

An der Untersuchung nahmen 35 Säuglinge im Alter von 3 bis 4 Monaten teil. Auch diese Kinder waren termgeboren und aus monolingualen Elternhäusern. Bei 11 Kindern musste das Experiment abgebrochen werden, so dass die endgültige Stichprobe wiederum aus 24 Kindern bestand. Das mittlere Alter der Gruppe betrug 3 Monate und 22 Tage bei einer Streubreite von 3 Monate 16 Tage bis 4 Monate 1 Tag. Bei 14 der Kinder handelte es sich um Mädchen, bei 10 um Jungen.

Untersuchungsmaterial

Es wurden dieselben Stimuli verwendet wie in Experiment 3.

Durchführung

Auch die Durchführung der Untersuchung entsprach im wesentlichen der Durchführung von Experiment 3. Da drei Monate alte Kinder jedoch noch Probleme bei spontanen Kopfdrehungen haben, wurde eine Veränderung in der Untersuchungskabine entsprechend des experimentellen Aufbaus von Bosch und Sebastián-Gallés (2001) für die Durchführung von Experimenten mit dem *Headturn Preference* Paradigma für jüngere Kinder vorgenommen. Die Lampen und Lautsprecher, die bei den älteren Kindern an den Seitenwänden der Kabine angebracht waren, wurden für die jüngeren Kinder rechts und links seitlich an der vorderen Kabinenwand angebracht. Zur Fixation der Seitenlampen war damit keine deutliche Kopfdrehung mehr notwendig. Die Fixierung konnte durch eine minimale Kopfdrehung oder allein durch eine Augenbewegung erfolgen.

Ergebnisse und Diskussion

Wiederum wurden die Orientierungszeiten für die einzelnen Durchgänge pro rhythmischem Muster für jedes Kind gemittelt. Für die trochäischen Zweisilbler zeigten

sich mittlere Orientierungszeiten von 12,72 s und für die jambischen Zweisilbler von 12,38 s. Dieser Unterschied war statistisch nicht signifikant ($t(df=23) = 0,58$; $p = 0,568$).

Dieses Muster deutet darauf hin, dass drei Monate alte Säuglinge im Gegensatz zu den knapp drei Monate älteren Kindern des vorherigen Experiments noch keine Präferenz für die trochäischen Muster zeigen. Dieser Befund spricht dafür, dass die Trochäuspräferenz tatsächlich eine Anpassung an die sprachlichen Gegebenheiten darstellt, die sich jedoch früher vollzieht, als nach den bisherigen Untersuchungsbefunden zu erwarten war.

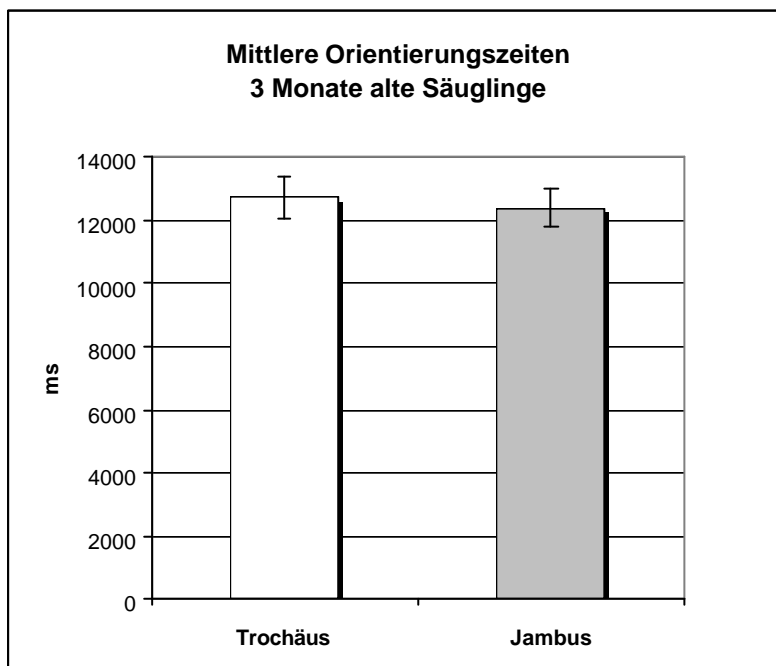


Abbildung 4: Mittlere Orientierungszeiten für trochäische und jambische Zweisilber aus Experiment 4

Zusammenfassende Diskussion

Die Befunde aus den Experimenten 3 und 4 stützen die Annahme einer sehr frühen Anpassung an die rhythmischen Verhältnisse in der Zielsprache. Deutschsprachige Säuglinge im Alter von sechs Monaten präferieren deutlich das trochäische Muster gegenüber dem jambischen. Da drei Monate alte Kinder diese spontane Präferenz noch

nicht zeigen, ist davon auszugehen, dass diese Präferenz nicht von Geburt an nachweisbar ist, sondern dass sie sich durch den zielsprachlichen Input entwickelt.

Es bleibt zu fragen, warum sich bei den sechs Monate alten englischsprachigen Säuglingen der oben dargestellten Studien noch keine Anzeichen einer Präferenz für den Trochäus fanden. Eine Möglichkeit könnten Unterschiede in der Verteilung der beiden Betonungsmustern im Englischen und Deutsch sein. Die gefundenen hohen Übereinstimmungen im Anteil der Inhaltswörter beider Sprachen, die mit einer starken Silbe beginnen, deuten aber eher auf relativ geringe Unterschiede hin, so dass diese Erklärung eher unplausibel erscheint.

Wahrscheinlich sind es eher Unterschiede in den experimentellen Vorgehensweisen, die die divergierenden Befunde ausmachen. Das Untersuchungsmaterial der Experimente zum Deutschen waren sehr einfache CV-CV-Zweisilbler, die sich lediglich im Rhythmus unterschieden und im Experiment sehr häufig wiederholt wurden. Dies könnte zu einer starken Reaktion der Säuglinge gerade auf die rhythmischen Merkmale geführt haben. In der im Vorgehen am besten vergleichbaren Studie von Jusczyk, Cutler und Redanz (1993) wurden dagegen als Stimulusmaterial je Rhythmustyp 96 englische Wörter verwendet, die sowohl im Segmentinventar als auch in den vorkommenden Silbenstrukturen sehr heterogen waren. Diese hohe Variabilität im Stimulusmaterial könnte dazu geführt haben, dass die sechs Monate alten Kinder dieser Studie nicht auf die rhythmischen Unterschiede reagiert haben.

Um die Frage nach der Rolle der Zielsprache bei der beobachteten Trochäuspräferenz endgültig zu klären, wäre jedoch die Durchführung vollkommen identischer Untersuchungen in verschiedenen Sprachen, die sich in Bezug auf die vorherrschende Rhythmuspräferenz deutlich unterscheiden, vonnöten. Hier würde sich beispielsweise ein Vergleich zwischen dem Deutschen und dem Französischen, das eine Vielzahl von endbetonten Inhaltswörtern aufweist, anbieten. Untersuchungen in diese Richtung fehlen jedoch bislang.

4.2.3 Prosodische Information und Wortkategorisierung

In Bezug auf die Frage, ob Säuglinge prosodische Information für die Kategorisierung von Wortklassen benutzen ist bislang so gut wie nichts bekannt. Einzig die Befunde von Shi, Werker und Morgan (1999) lassen erste Schlüsse darauf zu, dass bereits Neugeborene anhand von phonologischer Information eine Kategoriegrenze zwischen Elementen der offenen und der geschlossenen Klasse ziehen. In ihrem Habituationsexperiment wurden die Kinder zunächst mit typischen Mitgliedern der geschlossenen Klasse bzw. der offenen Klasse habituiert. Nach Erreichen des Habituationkriteriums wurde bei den Experimentalgruppen zu Vertretern der jeweils anderen Klasse gewechselt. Bei beiden Gruppen zeigte sich mit diesem Wechsel ein Anstieg der Saugraten.

Dieser Befund deutet darauf hin, dass bereits Neugeborene die typischen Merkmale von Elementen der geschlossenen Klasse von den typischen Merkmalen der Elemente der offenen Klasse differenzieren können. Da die im Experiment verwendeten Wörter aus kontinuierlicher Sprache herausgeschnitten worden waren, ist davon auszugehen, dass sich die Stimuli der beiden Klassen auch in prosodisch relevanten Merkmalen wie Dauer, Vokalqualität usw. unterschieden. Allerdings waren auch andere Merkmale, wie beispielsweise die Art des initialen Lautes und der phonotaktische Aufbau zwischen den beiden Klassen verschieden, so dass aus den Daten nicht ersehen werden kann, ob die Kinder speziell auf bestimmte Hinweise reagierten. Da sowohl während der Habituationsphase als auch während der Testphase verschiedene Exemplare der beiden Klassen präsentiert wurden, zeigen die Ergebnisse aber, dass bereits wenige Tage alte Säuglinge in der Lage sind, die gemeinsamen phonologischen Eigenschaften aus den präsentierten Token der beiden Klassen zu extrahieren und denen der jeweils anderen Klasse gegenüberzustellen. Damit ist eine wesentliche Voraussetzung zur Nutzung phonologischer Information zur Klassenbildung gegeben. Klare Evidenz, dass Säuglinge prosodische Muster zur Wortkategorisierung benutzen, liegt jedoch bislang nicht vor.

4.2.4 Prosodische Information und der Erwerb syntaktischen Wissens

4.2.4.1 Prosodische Information und Phrasensegmentierung

Wie oben beschrieben, werden syntaktisch relevante Grenzen häufig von prosodischen Grenzen begleitet, wobei diese dann auch häufig bestimmte phonetische Grenzmarkierungen wie spezifische Grundfrequenzverläufe, Dauer- und Intensitätsmerkmale aufweisen. Die Sensitivität gegenüber diesen Merkmalen könnte dazu beitragen, dass Kinder ihren sprachlichen Input überzufällig häufig in Einheiten segmentieren, die auch syntaktisch Einheiten darstellen. Dieses Herausfiltern syntaktisch relevanter Einheiten ist eine wesentliche Voraussetzung zum Erkennen der syntaktischen Regularitäten, die einer Sprache zugrundeliegen. Selbst wenn beispielsweise das Wissen über die Strukturabhängigkeit syntaktischer Regeln zum angeborenen Repertoire des Kindes gehören sollten, muss das Kind doch über Mittel verfügen die relevanten Struktureinheiten in seinem Input zu identifizieren. Durch die hohe Korrelation zwischen prosodischer und syntaktischer Struktur könnten hier also prosodische Merkmale wichtige Hilfsdienste leisten.

Eine der ersten Untersuchungen, die sich der Frage widmete, ob Kinder neben einer Sensitivität für prosodische Gliederungsmerkmale, diese Merkmale in irgendeiner Weise für ihre Sprachverarbeitung nutzen, stammt von Mandel, Jusczyk und Kemler Nelson (1994). Sie verfolgten die Hypothese, dass – wenn prosodische Information für die Strukturierung des Signals genutzt wird – prosodisch gegliederte Sätze von Kindern besser erinnert werden müssten als prosodisch nicht gegliederte. Dass Prosodie in dieser Weise die Verarbeitung gesprochener Sprache unterstützt, zeigt sich zum Beispiel darin, dass sich bessere Nachsprech- und Verstehensleistungen bei prosodisch gegliederten Sätzen gegenüber prosodisch oberhalb der Wortbetonung nicht strukturierten Wortketten zeigen (s. Weinert, 1996). In ihrem Experiment mit der *High-Amplitude Sucking* Methode fanden Mandel und Mitarbeiter, dass 2 Monate alte Säuglinge stärker auf die Veränderung von Lautsegmenten in einem Satz reagierten, wenn dieser in seiner gesamten Form eingesprochen worden war, als wenn er aus einzeln eingesprochenen Wörter synthetisiert worden war oder aus zwei Fragmenten zweier aufeinanderfolgend eingesprochenen Sätze zusammengeschnitten worden war (*cats like park benches* aus

Brigid really knows what cats like. Park benches are their favorite things to climb on). Insgesamt zeigt sich also, dass die Kinder Veränderungen im prosodisch wohlgeformten Satz offensichtlich besser bemerkten als im prosodisch gar nicht oder nicht adäquat gegliederten Satz. Allerdings bleibt dieser Effekt mit den unterschiedlichsten Erklärungen kompatibel: es könnte sein, dass die Kinder dem prosodisch wohlgeformten Sätzen mehr Aufmerksamkeit zuwenden als nicht wohlgeformten Sequenzen; es ist auch möglich, dass die prosodische Wohlgeformtheit zu besseren Gedächtnisleistungen führt. Die Ergebnisse lassen aber keinen direkten Schluss darauf zu, dass die Kinder prosodische Information zur Segmentierung von Sätzen heranziehen.

Es finden sich in der Arbeit von Mandel und Mitarbeitern aber auch Hinweise darauf, dass Säuglinge Veränderungen innerhalb einer einzigen prosodischen Einheit stärker bemerken, als Veränderungen die gleichzeitig zwei verschiedene prosodische Einheiten betreffen. Die Kinder reagierten auf die Veränderungen der Wortfolge in Sätzen (z. B. *Cats would jump benches* vs. *Cats jump wood benches*) stärker wenn die Sätze als Einheit eingesprochen worden waren, als wenn ihnen Sätze präsentiert wurden, die aus zwei einzeln produzierten Fragmenten zusammengeschnitten worden waren. Auch dieser Bund deutet darauf hin, dass eine kohärente prosodische Struktur die Gedächtnisleistungen verbessert.

Gezielter testete eine Untersuchung von Nazzi, Kemler Nelson, Jusczyk und Jusczyk (2000) die Frage, ob prosodische Information von Säuglingen tatsächlich für die Segmentierung gesprochener Sprache herangezogen wird. Das Stimulusmaterial für diese Untersuchungen bestand zunächst aus zwei Textpassagen, die medial die gleiche Sequenz von sieben Wörtern enthielten, wobei diese sieben Wörtern in den beiden Passagen jedoch syntaktisch unterschiedlich strukturiert waren, d. h. in einer Textpassage gehörten sie zu zwei aufeinanderfolgende Sätze – waren also durch eine Satzgrenze getrennt – während sie in der zweiten Textpassage einen gesamten Satz bildeten:

- (1) John doesn't know what rabbits eat. Leafy vegetables taste so good. They don't cost much either.
- (2) Many animals prefer some things. Rabbits eat leafy vegetables. Taste so good is rarely encountered.

Man beachte, dass beide Textpassagen die Sequenzen *Rabbits eat leafy vegetables* und *Leafy vegetables taste so good* enthalten, wobei erste nur in der zweiten Passage auch einem gesamten Satz und somit einer prosodischen Einheit entspricht und die zweite Sequenz nur in der ersten Textpassage als prosodische Einheit auftritt. Das Experiment umfasste eine Familiarisierungsphase, in der den Kindern eine der beiden 7-Wort-Sequenzen in beiden prosodischen Strukturierungen präsentiert wurden (d. h. z. B. *rabbits eat # leafy vegetables* herausgeschnitten aus Textpassage (1) und *# rabbits eat leafy vegetables #* herausgeschnitten aus Textpassage (2)). In der Testphase des Experiments wurden die beiden vollständigen Textpassagen je viermal präsentiert. Die 6 Monate alten Probanden zeigten jeweils längere Orientierungszeiten zu den Textpassagen, die die familiarisierte Wortsequenz als eine prosodische Einheit enthielt, gegenüber den Textpassagen, in denen die Wortsequenz in zwei Sätzen auftrat. Der gefundene Unterschied könnte darin begründet sein, dass die Detektion der familiarisierten Wortsequenz im Text leichter fällt, wenn diese dort eine prosodische Einheit darstellt. Dies wiederum würde dafür sprechen, dass die prosodische Information tatsächlich *on-line* während der Verarbeitung der Textpassagen zu deren Segmentierung und damit zur Strukturierung der Inputinformation herangezogen wird.

4.2.4.2 Prosodische Information und der Erwerb von Wortstellungsregeln

Die prosodische Struktur einer Äußerung korreliert nicht nur mit der syntaktischen Struktur durch die Beziehungen zwischen prosodischen und syntaktischen Grenzen, sondern offensichtlich korreliert die prosodische Struktur auch mit Wortstellungsregularitäten in einer Sprache (Mazuka, 1996; Nespor, Guasti & Christophe, 1996). Wie bereits in Kapitel 1 erwähnt lassen sich Beziehungen zwischen der Kopf-Komplement Stellung und der rhythmischen Struktur der phonologischen

Phrase in einer Sprache feststellen: in kopffinalen Sprachen wie dem Italienischen oder dem Französischen, in denen das Komplement dem Kopf folgt, tragen phonologische Phrasen ihre Hauptbetonung stets auf dem letzten Element. Demgegenüber folgt der Kopf in kopffinalen Sprachen wie dem Türkischen oder dem Japanischen seinem Komplement. In diesen Sprachen zeichnen sich phonologische Phrasen dadurch aus, dass sie die Hauptbetonung im ersten Element Demnach findet sich in kopffinalen Sprachen in verzweigenden phonologischen Phrasen stets ein schwach-stark-Muster, während sich in kopffinalen Sprachen in verzweigenden phonologischen Phrasen ein stark-schwach Muster findet.

Nespor und Mitarbeiter (Nespor et al., 1996; Guasti, Christophe, van Ooyen & Nespor, 2001) schlagen nun vor, dass für das sprachlernende Kind die rhythmische Struktur innerhalb der phonologischen Phrase die kritische Information im sprachlichen Input darstellt, die ihm hilft herauszufinden, ob seine Sprache kopffinal oder kopffinal ist und damit eine Grundlage für das Erlernen diverser einzelner Wortstellungsregeln der Zielsprache, die mit der Kopfrichtung verbunden sind (z. B. Verb-Objekt-Abfolge, Nomen-Adjektiv-Abfolge in der Nominalphrase etc.), bietet. Genauer wird dies im sogenannten *Rhythmic Activation Principle* im Rahmen des Parameter-Fixierungsmodells formuliert:

When you hear sequences of (ws) within an Intonational Phrase, set the Head-Complement parameter with the value Head-Complement. When you hear sequences of (sw) within an Intonational Phrase set the Head-Complement parameter with the value Complement-Head.

Guasti et al., 2001, S. 237

Dieses Prinzip wird als angeboren angenommen, so dass es für das Kind ausreicht, die entsprechende Input-Information zu identifizieren, um den Parameter korrekt zu fixieren. Dass Kinder diese Art von Input-Information relativ früh ausnutzen können, ist vor dem Hintergrund der vorliegenden empirischen Befunde sehr wahrscheinlich. Im wesentlichen sind zwei Arten von Fähigkeiten hierfür relevant: zum einen die Fähigkeit, zwischen verschiedenen rhythmischen Mustern zu unterscheiden. Auf der Ebene von Zweisilblern scheint nach den oben berichteten Befunden hieran keine Zweifel angebracht (Jusczyk & Thompson, 1978). Um die Prominenzverhältnisse innerhalb einer phonologischen Phrase zu bestimmen, ist allerdings eine Berechnung der

Betonungsstärken in einer größeren Domäne notwendig. Befunde von Guasti et al. (2001) deuten allerdings darauf hin, dass Säuglinge auch dies bereits in sehr jungem Alter können. Sie benutzten in einer Untersuchung resynthetisierte Sätze des kopffinalen Französisch und des kopffinalen Türkisch, in denen alle Vokale zum Schwa verändert worden waren, alle Verschlusslaute zu [p], alle Frikative zu [s], alle Nasale zu [m], alle Liquide zu [l] und alle Semivokale zu [j]. Unverändert blieb dagegen die prosodische Struktur der Äußerungen. In einer Untersuchung mit dem *High Amplitude Sucking* Verfahren, in denen zwei Monate alten Säuglingen diese segmental nicht mehr unterscheidbaren aber prosodisch verschiedenen Stimuli präsentiert wurden, fanden sich Hinweise, dass die Säuglinge die Stimuli aus den unterschiedlichen Sprachen diskriminieren können. Dies spricht dafür, dass Säuglinge tatsächlich sensitiv für die Prominenzverhältnisse innerhalb der phonologischen Phrase sind.

Für die Anwendbarkeit des *Rhythmic Activation Principle* ist es außerdem notwendig, dass die Kinder die für die rhythmische Strukturierung auf dieser Ebene relevanten prosodischen Einheiten - sprich die Intonationsphrasen und die phonologischen Phrasen - als Einheiten identifizieren können. Nach den oben berichteten Befunden zur Sensitivität gegenüber prosodischen Markierungen von Satz- und Phrasengrenzen (Hirsh-Pasek et al., 1987, Jusczyk et al., 1992) sollte dieses für Kinder ab dem 9. Lebensmonat kein Problem mehr darstellen.

Damit scheinen Säuglinge die perzeptuellen Voraussetzungen dafür zu besitzen, die prosodischen Verhältnisse in ihrer Zielsprache für das Erlernen von Wortstellungsregularitäten zu nutzen. Betrachtet man die Sensitivität für die dafür relevante prosodische Information, könnte man davon ausgehen, dass der entsprechende Parameter bereits gegen Ende des ersten Lebensjahres fixiert sein könnte. Damit vereinbar sind Befunde, nach denen Wortstellungsfehler in den ersten Produktionen so gut wie nicht auftreten, d. h. sobald Kinder beginnen Mehrwortäußerungen zu produzieren, entsprechen diese im wesentlichen den Wortstellungsregularitäten der Zielsprache (Clahsen, 1982; Penner & Weissenborn, 1996; Rothweiler, 1989; Weissenborn, 1991, 1994).

Ein Problem für diesen Ansatz stellen allerdings Sprachen dar, die keine einheitliche Kopfposition aufweisen, sondern wo diese variieren kann. Zu diesen Sprachen gehören das Niederländische und das Deutsche. Die Variabilität der

Kopfposition sei für das Deutsche anhand der verschiedenen Möglichkeiten zur Positionierung von Verb und Objekt innerhalb der Verbalphrase erläutert. Im Deutschen kann das finite Verb entweder final nach dem Objekt auftauchen, wie es immer in durch einen Komplementierer eingeleiteten Nebensätzen der Fall ist, es kann aber auch in zweiter Position vor dem Objekt auftreten:

(a) Fritz sagt, dass Oma Opa sucht

(b) Fritz sagt, Oma sucht Opa

In beiden Fällen erwarten wir, dass das Objekt als Komplement die höhere prosodische Prominenz aufweist. Demnach müssten also in Satz (a) die phonologische Phrase [Opa sucht] anfangsbetont sein, während in Satz (b) die phonologische Phrase [sucht Opa] endbetont sein müsste. Eigene prosodische Analysen deutscher Sätze mit Verbend- und mit Verbzweitstellung zeigen, dass die prosodischen Prominenzverhältnisse in Sätzen mit Objekt-Verb Folge und in Sätzen mit Verb-Objekt Folge tatsächlich variieren (Höhle, Weissenborn, Schmitz & Ischebeck, 2001).

Die Frage ist nun aber, in welcher Weise Kinder mit dieser Nonkonformität der prosodischen Struktur innerhalb der phonologischen Phrase umgehen. Aufgrund der widersprüchlichen Input-Information sollten sie Schwierigkeiten haben, den Kopfrichtungs-Parameter im Sinne des *Rhythmic Activation Principle* zu fixieren und damit Probleme im Erwerb der Wortstellungsregeln haben. Aber auch deutsche Kinder verfügen bereits im dritten Lebensjahr über Kenntnisse der Verbstellungsregeln ihrer Zielsprache und ihrer Abhängigkeit vom Satztyp (Höhle et al., 2001).

Guasti et al. (2001) schlagen vor, dass die enorme Sensitivität für die Frequenz von Mustern in der Input-Sprache zur Lösung des Variabilitätsproblems beitragen könnte. Hiernach könnten Kinder das frequentere rhythmische Muster in der phonologischen Phrase zur Grundlage der Parameterfixierung heranziehen. Die Autoren betonen zudem, dass im Niederländischen (wie auch im Deutschen) das die Verbendstellung charakterisierende stark-schwach Muster in der Verbalphrase stets am Satzende auftaucht, da das Verb satzfinal auftritt. Diese positionale Beschränkung könnte den Kinder einen zusätzlichen Hinweis auf die Besonderheit dieses Muster geben. Gezieltere Untersuchungen zum Zusammenhang von prosodischen Mustern und

dem Erwerb von Wortstellungsregularitäten fehlen zwar bislang, aber eigene Befunde aus einer Untersuchung zur Frage, ab wann deutschsprachige Kinder Sätze mit korrekter Verbstellung von Sätzen mit falscher Verbstellung unterscheiden können, deuten auf einen Einfluss prosodischer Information (s. Höhle et al., 2001).

In dieser Untersuchung mit dem *Headturn Preference* Paradigma fand sich bereits bei Säuglingen ab einem Alter von 18 Monaten eine Diskriminationsfähigkeit zwischen Nebensätzen mit korrekter und Nebensätzen mit falscher Verbstellung, wenn der Nebensatz ein Objekt enthielt (*Bert sagt, dass Lisa Oma hilft* vs. *Bert sagt, dass Oma hilft Lisa*). Dagegen unterschieden die Kinder nicht zwischen korrekter und falscher Verbstellung bei entsprechenden Sätzen, die statt des Objekts ein Adverb enthielten (*Bert sagt, dass Lisa draussen spielt* vs. *Bert sagt, dass Lisa spielt draussen*). Die akustische Analyse der Prominenzverhältnisse der jeweils letzten beiden Wörter dieser Sätze, zeigte deutliche prosodische Unterschiede zwischen der grammatischen und der entsprechenden ungrammatischen Version der Sätze mit Objekt, während sich die grammatischen und die ungrammatischen Versionen der Sätze mit Adverb prosodisch nicht so deutlich unterschieden. Diese Befunde der akustischen Analyse stützen die Annahme systematischer Beziehungen zwischen Wortstellungsregularitäten und prosodischen Prominenzverhältnissen. Dies könnte zudem bedeuten, dass die für die Kinder gefundene Diskriminationsfähigkeit zwischen den grammatischen und den ungrammatischen Sätzen mit Objekt auf den deutlichen prosodischen Unterschieden zwischen den grammatischen und ungrammatischen Sätzen beruht. Dies wäre als weitere Evidenz dafür zu werten, dass die Kinder in der Lage sind, Prominenzverhältnisse innerhalb größerer Domänen zu berechnen, direkte Evidenz dafür, dass die prosodische Information für den Erwerb von Wortstellungsregularitäten herangezogen wird, gibt es bislang jedoch nicht.

4.3 Möglichkeiten und Grenzen des Prosodischen Bootstrappings

In den vorangehenden Abschnitten wurde gezeigt, dass Säuglinge eine hohe Sensitivität gegenüber den diversen prosodisch-akustischen Eigenschaften des sprachlichen Signals haben, die in systematischer Weise mit der linguistischen Struktur der sprachlichen Äußerung korrelieren, und dass sie diese Eigenschaften offensichtlich auch zur Strukturierung des sprachlichen Signals heranziehen. Diese Befunde unterstützen in ihrer Gesamtheit die Annahme, dass für die frühe Sprachverarbeitung und für den frühen Spracherwerb prosodisch-rhythmische Eigenschaften des sprachlichen Inputs eine entscheidende Rolle spielen. Somit bestätigt sich ein wesentlicher Teil des Ansatzes des „prosodischen Einstiegs“ in den Syntaxerwerb.

Zwei Punkte dieses Ansatzes bedürfen jedoch weiterer Diskussion. Erstens muss innerhalb dieses angenommen werden, dass die akustischen Hinweisreize, die das Kind heranzieht, ohne bereits vorhandenes sprachliches Wissen, also in einer reinen datengetriebenen Weise verarbeitet und interpretiert werden können. Zweitens muss angenommen werden, dass die Hinweisreize eine hohe Reliabilität in Hinblick auf die zugrundeliegende Struktur der Äußerung haben, d. h. dass das Vorhandensein bzw. Fehlen eines bestimmten akustischen Reizes in eindeutiger Weise auf eine andere linguistische Strukturebene abgebildet werden kann. Beide Annahmen werden im folgenden näher beleuchtet.

4.3.1 Können akustische prosodische Markierungen rein datengetrieben verarbeitet werden?

4.3.1.1 Multifunktionalität akustisch-prosodischer Merkmale

Ein wesentliches Problem der Verarbeitung prosodischer Information, vor allen Dingen der Interpretation prosodischer Merkmale, ist die Tatsache, dass die phonetischen Korrelate ihrer Markierung für verschiedenste Funktionen eingesetzt werden. Wie gezeigt, wird beispielsweise die Vokaldauer als eines der wesentlichen akustischen Korrelate von Betonung angenommen. Die Länge eines Vokals ist jedoch nicht nur von

diesem Parameter abhängig sondern auch von der Position einer Silbe in der syntaktischen bzw. prosodischen Struktur einer Äußerung, da die Vokallänge auch das akustische Korrelat der finalen Silbenlänge darstellt (z.B. Klatt, 1975, 1976). Darüber hinaus hängt die Länge eines Vokals auch von seiner Qualität und dem segmentalen Kontext innerhalb eines Wortes ab. So ist ein Vokal beispielsweise meist länger, wenn ihm ein stimmhafter Konsonant folgt, als wenn ihm ein stimmloser Konsonant folgt (Crystal & House, 1988). Weiter kommt erschwerend hinzu, dass in vielen Sprachen – wie auch dem Deutschen – die Vokallänge phonemunterscheidend ist. Die Multifunktionalität des Dauermerkmals könnte also zum Beispiel dazu führen, dass ein inhärent kurzer Vokal in bestimmten Positionen, beispielsweise am Ende einer Phrase, die gleiche Länge aufweist wie ein inhärent langer Vokal innerhalb einer Phrase. Damit stellt sich genau das Zirkularitäts-Problem, für das die Bootstrapping-Ansätze eine Lösung anbieten sollten: die Feststellung der absoluten Dauer eines Vokals reicht nicht aus, um entscheiden zu können, ob dieser Vokal Teil einer betonten Silbe oder einer unbetonten Silbe ist, und sie reicht auch nicht aus um zu entscheiden, ob die Vokaldauer eine folgende Strukturgrenze markiert. Vielmehr muss das Verarbeitungssystem – zumindest in Sprachen, in der Dauer ein phonemunterscheidendes Merkmal ist - feststellen, ob der Vokal in einer betonten Silbe auftritt, ob die kritische Silbe vor einer relevanten Grenze auftritt und welcher Art der umgebende Kontext ist, bevor das akustische Ereignis einer bestimmten Phonemkategorie zugeordnet werden kann. Umgekehrt kann eigentlich aber erst festgestellt werden, ob ein Vokal als Kennzeichen einer folgenden strukturellen Grenze gelangt ist, wenn deutlich ist, ob in dem Wort, in dem der Vokal enthalten ist, ein inhärent kurzer oder ein inhärent langer Vokal auftreten sollte. Diese Überlegungen deuten daraufhin, dass einzelne phonetische Merkmale keine zuverlässigen Indikatoren prosodischer Struktur sind. Die Frage ist, ob dies auch für die anderen phonetischen Korrelate prosodischer Struktur gilt.

Für die Grundfrequenz muss diese Frage ebenfalls positiv beantwortet werden. Wie gesagt, haben betonte Silben im allgemeinen eine höhere Grundfrequenz als unbetonte Silben. Am Ende einer prosodischen Einheit finden sich häufig charakteristische Bewegungen der Tonhöhe, die sogenannten Grenztöne. Aber sowohl die Grundfrequenzwerte einer einzelnen Silbe als auch die Intonationskontour innerhalb einer Äußerung wird von einer Reihe unterschiedlicher Faktoren bestimmt (s. z. B.

Fisher & Tokura, 1996). So ist die Intonationskontur einer Äußerung massiv von pragmatischen und informationsstrukturellen Aspekten beeinflusst. So kann der Punkt der höchsten Tonhöhe in einem Satz bei gleicher Wortfolge an den verschiedensten Positionen sein, je nachdem wo der Fokus der Äußerung liegt. Überlagert wird die Fokusintonation noch davon, welche pragmatische Funktion der Satz erfüllt. So hat beispielsweise ein Fragesatz eine deutlich andere Intonationskontur als ein Aussagesatz.

Auch in Bezug auf die Grundfrequenz ist also nicht davon auszugehen, dass ein bestimmter Wert absolut interpretierbar ist, sondern lediglich vor dem Hintergrund des Einbeziehens größerer Domänen. Bislang ist weitgehend ungeklärt, auf welche Weise das Verarbeitungssystem diese verschiedenen Größen, die den Intonationsverlauf einer Äußerung bestimmen, differenziert berechnen kann.

Es zeigt sich also, dass sowohl die Dauer als auch die Tonhöhe von einzelnen Silben im akustischen Reiz Informationen verschiedener sprachlicher Ebenen tragen. Die konkreten Werte, die die akustischen Parameter in einer Äußerung annehmen, sind das Resultat einer Interaktion von segmentalen, rhythmischen, syntaktischen und informationsstrukturellen Merkmalen. Daher können die Werte einzelner akustischer Parameter auch nicht direkt in Hinblick auf eine sprachliche Ebene, z. B. als Grenzmarkierung für eine Phrase interpretiert werden, sondern nur vor dem Hintergrund aller Faktoren, die ihren Wert bestimmen. Genau dieses kann jedoch eine rein datengetriebene *bottom-up* Verarbeitung nicht leisten.

4.3.1.2 Sprachspezifische Aspekte prosodischer Markierungen und prosodischer Struktur

Das Problem der Sprachspezifik ist ein weiteres, dass eine ausschließlich vom sprachlichen Signal bestimmte Analyse sprachlicher Merkmale und deren Abbildung auf höhere sprachliche Ebenen während des Spracherwerbs unplausibel erscheinen lässt. Einzelsprachspezifische Merkmale bzw. einzelsprachspezifische Korrelationen zwischen akustischen und strukturellen Eigenschaften sind prinzipiell schlechte Kandidaten für Bootstrapping-Mechanismen, da ihre Erkennung selbst bzw. ihre Ausnutzung bereits Lernprozesse in Richtung auf spezifische Merkmale der Muttersprache voraussetzt.

Einzel sprachliche Ausprägungen sind für den Ansatz des prosodischen Einstiegs in zweierlei Hinsicht relevant. Zum einen können strukturell relevante prosodische Merkmale in verschiedenen Sprachen unterschiedlich phonetisch-akustisch realisiert sein. So variieren die Korrelate für Wortbetonung in verschiedenen Sprachen erheblich. Während in einer Reihe von Sprachen die Länge des Silbenkerns für den Betonungsstatus die ausschlaggebende Rolle spielt, wie es beispielsweise im Deutschen der Fall ist, ist in anderen Sprachen, wie beispielsweise dem Französischen, die Tonhöhe das entscheidende akustische Merkmal (Jessen et al., 1985; Meisenburg & Selig, 1998). Geht man davon aus, dass alle Korrelate der Betonung dazu führen, dass betonte Silben akustisch hervorgehobener sind als unbetonte, mag die Sprachspezifik in diesem Bereich für das Kind kein großes Problem darzustellen. Wenn das Wahrnehmungssystem quasi automatisch zwischen salienten und weniger salienten Silben unterscheiden kann, dann kann rein deduktiv auf die Eigenschaft, die diese akustische Differenz ausmacht, geschlossen werden.

Schwieriger stellt sich die Herstellung der Korrelation zwischen einer akustischen Eigenschaft und ihrer sprachspezifischen Funktion dar. So können gleiche akustische Eigenschaften in unterschiedlichen Sprachen ganz verschiedene Funktionen haben. Beispielsweise hat im Englischen der Tonhöhen-Akzent im wesentlichen die Aufgabe der Hervorhebung einzelner Silben für die Markierung von fokussierten Einheiten. Im Koreanischen und Japanischen hat dasselbe akustische Merkmal jedoch eine vollkommen andere Funktion: hier dient der Tonhöhen-Akzent zur Markierung der Grenze einer prosodischen Einheit (Venditti, Jun & Beckman, 1996). Auf welche Art kann das Kind entscheiden, welche Funktion dieses akustische Merkmal in seiner Umgebungssprache hat?

Ein weiteres Problem stellt die Tatsache dar, dass Sprachen sich dadurch unterscheiden können, ob ein bestimmtes akustisches Merkmal den Anfang oder das Ende einer strukturellen Einheit markiert. Da die akustischen Korrelate prosodischer Information, mit Ausnahme von Pausen nicht direkt in der Grenze, sondern in den adjazenten sprachlichen Elementen zu finden sind, kann ein wesentliches Merkmal im Prinzip der Grenze vorangehen oder ihr folgen. Ein einfaches Beispiel dafür ist die schon im Detail behandelte Wortbetonung. Wie gesehen dominiert im Englischen das trochäische Betonungsmuster und englische Kinder nutzen diese Dominanzverhältnisse

sehr früh für die Segmentierung der gesprochenen Sprache aus, indem sie Trochäen als strukturelle Einheiten und potentielle Wörter behandeln. In anderen Sprachen, wie beispielsweise dem Französischen, dominiert jedoch im lexikalischen Inventar die Endbetonung. Würde dagegen ein französisches Kind, in dessen Sprache das trochäische Muster nicht vorherrschend ist, von der gleichen Hypothese ausgehen und denselben Segmentierungsmechanismus benutzen, müsste es bei einer Vielzahl der Wörter, die es hört, zunächst zu Fehlsegmentierungen gelangen. Wie erwähnt, ist nichts darüber bekannt, ob französische Säuglinge zunächst nach einer solchen Strategie segmentieren. Diese Annahme ist allerdings recht unplausibel, da man in diesem Fall davon ausgehen müsste, dass der Lexikonerwerb des französischen Kindes generell schlechtere Startbedingungen hat als der Lexikonerwerb des englischen Kindes. Wenn dies so wäre, wären sicherlich deutlichere Unterschiede im Lexikonerwerb zwischen französischen und englischen Kindern beobachtet worden, die jedoch meines Wissens nicht beobachtet wurden.

Es ist also zu vermuten, dass der sprachliche Input Informationen enthält, die es den Kindern sehr früh erlaubt zu entscheiden, ob in ihrer Sprache betonte Silben eher wortinitial oder wortfinal auftauchen. Jusczyk et al. (1999) haben die Vermutung geäußert, dass in diesem Zusammenhang Eigennamen und Diminutivformen eine entscheidende Rolle spielen. Im Englischen zeigt sich bei Eigennamen und Diminutivformen eine sehr deutliche Dominanz des Trochäus, die noch größer wird, wenn man Koseformen in die Betrachtung einbezieht: auch die Koseformen von eigentlich nicht trochäischen Namen stellen zum Großteil trochäische Zweisilbler dar, wie z. B. *Chrissy* für *Christine* und *Chelly* für *Michele*. Nun stellt der eigene Name sicherlich eines der frequentesten Wörter im sprachlichen Input des Kindes dar, den Kinder auch bereits im Alter von 4 bis 5 Monaten erkennen können (Mandel, Jusczyk & Pisoni, 1995). Möglicherweise bildet also die Erkennung der prosodischen Verhältnisse innerhalb dieser für das Kind hochfrequenten Form den Einstieg in die Wortsegmentierung.

Auch in Bezug auf die oben erwähnten Tonhöhen-Akzente als Markierungen von Grenzen prosodischer Phrasen können sich sogar innerhalb verschiedener Dialekte einer Sprache Unterschiede in Hinblick auf die Position des Akzents in Relation zur Grenze zeigen. So erscheint im Standard-Seoul-Koreanischen der Akzent am Ende der

prosodischen Phrase während er in einem anderen Dialekt des Koreanischen (Chonnam) auf den initialen Elementen der Phrase realisiert wird (Venditti et al., 1996). Isoliert kann also der Tonhöhen Akzent trotz der hohen Korrelation seines Auftretens mit einer prosodischen Grenze keine Auskunft über die genaue Position dieser Grenze geben. Nach dieser Beobachtung schätzen Venditti und Mitarbeiter den Ansatz des Prosodischen Einstiegs denn auch eher pessimistisch ein:

However, the differences among the languages also show that the prosodic categories and their mappings to other linguistic structures are arbitrary enough that we cannot guess a priori that prosody will play a more important role in acquiring the syntactic structures than syntactic structures will play in acquiring the prosodic structures.

Venditti et al., 1996, S. 311

4.3.2 Reliabilität prosodischer Merkmale als Grenzmarkierung

4.3.2.1 Die Zuverlässigkeit akustischer Grenzmarkierungen

Eine weitere Frage, die in Hinblick auf die Plausibilität des *Prosodic-Bootstrapping* Ansatzes betrachtet werden muss, ist die Frage nach der Zuverlässigkeit prosodischer Merkmale als Hinweis auf das Vorliegen einer syntaktischen Grenze. Bei der Bestimmung dieser Zuverlässigkeit müssen zwei Aspekte getrennt voneinander betrachtet werden: erstens der Grad der Wahrscheinlichkeit, dass mit dem Auftreten eines bestimmten akustischen Merkmals oder eines Bündels akustischer Merkmale, das das Vorhandensein einer prosodischen Grenze signalisiert, auch eine syntaktische Grenze einhergeht und zweitens der Grad der Wahrscheinlichkeit, dass eine syntaktische Grenze auch tatsächlich prosodisch markiert ist. Fernald und McRoberts (1996) argumentieren dafür, dass für das sprachlernende Kind vor allen Dingen der erste Aspekt von entscheidender Bedeutung ist, da das Kind durch die Wahrnehmung eines bestimmten akustischen Merkmals auf das Vorhandensein bestimmter struktureller Merkmale schließt und so dem Gehörten eine Struktur zuweist. Das Auftauchen eines in

dieser Hinsicht relevanten akustischen Merkmals an einer strukturell nicht relevanten Position im Sinne eines falschen Alarms müsste zwangsläufig zu einer falschen Strukturzuweisung führen und somit zum Aufbau einer falschen Strukturrepräsentation führen. Demgegenüber führt das Fehlen einer akustischen Markierung an einer syntaktischen Grenze nicht unmittelbar zu einer fehlerhaften Analyse des sprachlichen Inputs, sondern es fehlt lediglich die unterstützende Wirkung dieser Information.

Das mögliche Vorhandensein von Fällen falschen Alarms im Input, d. h. des Auftretens von Grenzsignalen an Stellen, an denen keine syntaktische Grenze ist, ist natürlich wiederum in erster Linie durch die Multifunktionalität der akustischen Parameter bedingt. Betrachten wir zunächst den Grad der Korrelation des Auftretens einer Pause mit dem Vorliegen einer syntaktischen Grenze. Zahlreiche Untersuchungen bestätigen, dass an syntaktischen Grenzen tatsächlich mit hoher Wahrscheinlichkeit Pausen auftreten, wobei die Pausenlänge sogar die syntaktische Hierarchie reflektiert (Goldman-Eisler, 1972; Grosjean & Collins, 1979; Grosjean, Grosjean & Lane, 1979). Analysen, die der Frage nachgingen, wie häufig Pausen tatsächlich an syntaktischen Grenzen auftreten, zeigen allerdings, dass bis zu 50% aller Pausen innerhalb von syntaktischen Phrasen und nicht deren Grenzen auftreten (Grosjean et al., 1979; Hawkins, 1971; Maclay & Osgood, 1959).

Über die Reliabilität des Grundfrequenzverlaufs als Indikator für syntaktische Grenzen ist weniger bekannt als über Pausen. Wie oben schon gesagt, ist der Grundfrequenzverlauf innerhalb eines Satzes stark von informationsstrukturellen und pragmatischen Faktoren abhängig, so dass auch innerhalb eines Satzes oder Teilsatzes mit massiven Veränderungen im Grundfrequenzverlauf zu rechnen ist. Es ist daher zu vermuten, dass dieses Merkmal für sich genommen relativ unzuverlässig ist. Analoges gilt auch für die Segmentlänge, da diese – wie oben bereits ausgeführt – keinen bestimmten Wert als Indikator für eine prosodische Grenze annimmt, sondern von vielfältigen Faktoren bestimmt wird. Zudem findet sich das Phänomen der finalen Längung auf den unterschiedlichsten Ebenen der linguistischen Struktur: es betrifft satzfinale, phrasenfinale aber auch wortfinale Silben. Die Frage ist also, wie das Merkmal der Länge in seiner funktionalen Vielfältigkeit noch strukturell genutzt werden kann.

Ein grundsätzliches Problem betrifft den zweiten Aspekt der Zuverlässigkeit, nämlich die Frage nach dem tatsächlichen Vorhandensein dieser akustischen Merkmale an den strukturell relevanten Grenzen. Ein Großteil der Analysen prosodischer Grenzmarkierungen basiert auf Sprachkorpora aus eher formalen Sprechsituationen wie Fernsehdiskussionen oder Konferenzbeiträgen bzw. auf gelesenen Texten. Inwieweit die Ergebnisse für dieses Material auf „normale Spontansprache“ übertragen werden können, ist fraglich, da die Pausensetzung von Kompetenzfaktoren wie beispielsweise der Sprechgeschwindigkeit abhängig ist (Grosjean & Collins, 1979) und die beim lauten Lesen produzierte Sprache deutlichere prosodische Merkmale aufweist als Spontansprache (Lieberman, Katz, Jongman, Zimmerman & Miller, 1985; Umeda, 1982). Dieses muss im übrigen auch bei der Bewertung der Ergebnisse aller psycholinguistischen Untersuchungen, die einen Einfluss der prosodischen Struktur auf die Verarbeitungsprozesse bzw. Interpretation syntaktischer ambiger Sätze festgestellt haben, bedacht werden. In diesen Untersuchungen bestand das verwendete Stimulusmaterial stets aus für diesen Zweck eingesprochenen, gelesenen Sätzen. Die Untersuchungen können danach nur zeigen, dass Sprecher diese Information nutzen, wenn sie im Signal vorhanden ist; inwieweit das Signal diese Information jedoch in der normalen Kommunikationssituation tatsächlich enthält, ist noch weitgehend unklar.

Allerdings mag im Zusammenhang mit der Plausibilität eines prosodischen Einstiegs in den Syntaxerwerb die Standardsprache auch gar die nicht relevante Datengrundlage sein, sondern eher die Merkmale des kindgerichteten Sprachregisters. Analysen dieses Sprachregisters zeigen eine relativ hohe Korrelation zwischen prosodischen Merkmalen und strukturellen Grenzen sowie eine starke Salienz dieser prosodischen Merkmale. Im Vergleich zu an Erwachsene gerichtete Sprache ist die Längung finaler Segmente in kindgerichteter Sprache sehr deutlich (Bernstein Ratner, 1986), Pausen sind sehr lang (Fernald et al., 1989) und die Veränderungen der Tonhöhe vor Pausen sind sehr ausgeprägt (Stern, Spieker, Barnett & MacKain, 1983). Zudem ist die Realibilität dieser Merkmale als Hinweis auf eine syntaktische Grenze offensichtlich sehr hoch: in kindgerichteter Sprache tritt ein Großteil der Pausen tatsächlich an syntaktischen Grenzen auf (Dale 1974; Fernald & Simon, 1984). Demnach könnten die prosodischen Merkmale kindgerichteter Sprache tatsächlich – wie von vielen Autoren vorgeschlagen – dem Kind den Einstieg in die strukturelle Analyse seines Inputs

erleichtern. Fernald and McRoberts (1996) äußern sich jedoch aus einem anderen Grund skeptisch: sie weisen daraufhin, dass viele der prosodisch als Einheit gekennzeichneten Äußerungen in kindgerichteter Sprache syntaktisch unvollständig sind (Newport, 1977; Fernald et al., 1989), so dass eine Abbildung prosodischer Einheiten auf syntaktische Einheiten das Kind auf einen falschen syntaktischen Weg bringen würde. Diese Beobachtung leitet auf ein generelles Problem der Abbildung prosodischer Struktur auf die syntaktische Struktur über, nämlich auf die Tatsache, dass zwischen prosodischer und syntaktischer Struktur keine ein-eindeutige Beziehung besteht.

4.3.2.2 Die Eigenständigkeit von prosodischer und syntaktischer Struktur

In den vorangegangenen Abschnitten wurde davon ausgegangen, dass die prosodischen Einheiten einer Äußerung in hohem Maß mit den syntaktischen Einheiten einer Äußerung übereinstimmen. Nur in diesem Fall können die akustischen Merkmale, die die Grenzen prosodischer Einheiten markieren, auch als Markierungen der Grenzen syntaktischer Einheiten genutzt werden. Theoretische Ansätze in der Phonologie betonen jedoch, dass die prosodische Struktur nicht in Abhängigkeit von der Syntax besteht, sondern eine eigenständige Domäne darstellt (z.B. Nespor & Vogel, 1986; Selkirk, 1984).

Prosodische Einheiten leiten sich danach nicht direkt aus syntaktischen Einheiten ab, sondern sie ergeben sich aus phonologisch relevanten Domänen: ähnlich wie zur Beschreibung syntaktischer Regeln syntaktisch relevante Domänen wie Phrasen, Sätze und Teilsätze notwendig sind, beziehen sich phonologische Regeln auf phonologisch relevante Domänen wie das phonologische Wort, die phonologische Phrase oder die Intonationsphrase, die eben nicht deckungsgleich mit syntaktischen Einheiten sind. Zwar gibt es Bereiche relativ systematischer Beziehungen zwischen prosodischen und syntaktischen Einheiten, so sind Satzgrenzen meist auch Grenzen von Intonationsphrasen, die durch die oben beschriebenen Merkmale prosodisch recht gut markiert sind, aber die prosodische Gliederung innerhalb eines Satzes kann sehr variabel sein. So können optionale prosodische Grenzen innerhalb eines Satzes an fast allen Positionen auftreten. Das Auftreten dieser optionalen Grenzen in einer Äußerung ist dann auch abhängig von Performanzfaktoren. In dieser Hinsicht ist beispielsweise

die Länge eines Satzes bzw. seiner Konstituenten ausschlaggebend: je länger ein Satz ist, desto geringer ist die Wahrscheinlichkeit, dass er prosodisch aus einer einzigen Intonationsphrase besteht. So wird beispielsweise die Grenze zwischen Subjekt und Prädikat eines Satzes eher prosodisch markiert, wenn das Subjekt relativ komplex ist, wie im Fall *das Mädchen mit den blonden Haaren # verließ das Haus ihrer Mutter* gegenüber dem einfacheren Subjekt *das Mädchen verließ das Haus ihrer Mutter*, wo mit höherer Wahrscheinlichkeit keine prosodische Grenze nach dem Subjekt auftritt (vgl. Ferreira, 1993; Grosjean et al., 1979).

Die Optionalität der Positionierung prosodischer Grenzen kann zudem dazu führen, dass eine prosodische Grenze innerhalb einer syntaktischen Einheit auftritt. Nach Selkirk (1984) wird die Position einer optionalen prosodischen Grenze vor allen Dingen von der Informationsstruktur einer Äußerung bestimmt, die syntaktisch ein und demselben Satz diverse prosodische Strukturen zuweisen kann. So kann beispielsweise der Satz *Jane gave the book to Mary* mit 6 verschiedenen prosodischen Phrasierungen realisiert werden, je nachdem welche Teile des Satzes fokussiert werden:

- (a) #Jane gave the book to Mary#
- (b) #Jane # gave the book to Mary#
- (c) #Jane gave the book # to Mary#
- (d) #Jane gave # the book # to Mary#
- (e) #Jane # gave the book # to Mary#
- (f) #Jane # gave # the book # to Mary#

(aus Selkirk, 1984; S. 293)

Zwar befinden sich alle prosodische Grenzen an syntaktisch relevanten Punkten, aber die syntaktische Phrasenstruktur deckt sich mit der prosodischen Phrasierung ausschließlich im Beispiel (b).

Systematische Widersprüche zwischen syntaktischer oder lexikalischer Struktur einer Äußerung und ihrer prosodischen Struktur finden sich dann, wenn einzelne syntaktische Komponenten ausschließlich aus einem Funktionswort bestehen. Funktionswörter bilden als unbetonte Elemente keine separaten prosodischen Konstituenten (Selkirk, 1996). Dies führt zum einen dazu, dass beispielsweise ein

pronominales Subjekt nur unter der sehr markierten Bedingung der Fokussierung dieses Elements durch eine prosodische Grenze von der folgenden syntaktischen Phrase abgegrenzt würde. Zweitens kann die phonologische Insalienz dieser Elemente dazu führen, dass sie zusammen mit einer anderen lexikalischen Einheit durch Klitisierung ein einziges phonologisches Wort bilden. Beispielsweise ist im Englischen häufig die Klitisierung des Auxiliars *is* an die vorangehende Konstituente zu beobachten: *The woman's going to the party*. Hier läuft die prosodische Struktur nicht nur der syntaktischen Struktur entgegen, sondern auch den Grenzen lexikalischer Einheiten.

Wie erwähnt gehen Fernald und McRoberts (1996) davon aus, dass insbesondere die Fälle eines falschen Alarms, d. h. die Fälle, in denen Grenzmarkierungen auftreten, ohne dass eine relevante Grenze vorliegt, zu Fehlanalysen des Input führen sollten. In diesem Zusammenhang muss auch noch einmal die rhythmische Wortsegmentierung erwähnt werden, die auch systematische Bereiche der Fehlsegmentierung beinhaltet.

Trotz der Unzulänglichkeiten in der Beziehung zwischen prosodischer Information und Syntax sind Kinder in der Lage, die spezifischen syntaktischen Gesetzmäßigkeiten, wie beispielsweise Wortstellungsregularitäten, ihrer Muttersprache zu erwerben. Da diese Regeln im allgemeinen strukturbezogen sind, ist es eine notwendige Voraussetzung relativ schnell die Grenzen der syntaktischen Einheiten im sprachlichen Input identifizieren zu können. Nach den obigen Ausführungen ist hierzu eine rein prosodische Vorgehensweise nicht ausreichend. Da jedoch massive empirische Evidenz für die Nutzung prosodischer Information für die Strukturierung des sprachlichen Inputs vorliegt, ist die Annahme plausibel, dass die prosodische Information tatsächlich zumindest für einen ersten Einstieg in die strukturelle Analyse des Input genutzt wird, jedoch sehr schnell zusätzliche Informationsquellen hinzugezogen werden, so dass die mangelnde Isomorphie zwischen Prosodie und anderen Strukturebene kein tatsächliches Problem für das lernende Kind darstellt. Dass Kinder relativ früh das alleinige Vertrauen auf prosodische Information aufgeben, ist im Bereich der Wortsegmentierung gut belegt. Bereits die Tatsache, dass 11 Monate alte Kinder in der Lage sind, auch jambische Wörter im Kontext zu erkennen, zeigt, dass sie keine reine rhythmische Segmentierung mehr anwenden. Die Frage ist nun, welche anderen Informationsquellen die Kinder zu nutzen beginnen. In folgenden Kapitel wird argumentiert, dass die Erkennung und Verarbeitung häufig wiederkehrender

Segmentsequenzen im sprachlichen Input, d. h. in Sprachen wie dem Englischen und dem Deutschen vor allen Dingen Funktionswörter und Flexionsendungen mit ihrer typischen hohen Gebrauchshäufigkeit, in diesem Zusammenhang eine entscheidende Rolle spielt.

5 Segmentale und segmental-sequentielle Information als strukturelle Hinweise

5.1 Segmentale und sequentielle Information für strukturelle Grenzen

In diesem Kapitel steht die Frage im Vordergrund, welche Art oder Arten nichtprosodischer Information im Signal als Markierung struktureller Eigenschaften dienen können. Hier kommen zum einen akustisch-phonetische Merkmale einzelner Lautsegmente bzw. von Lautsequenzen in Frage sowie distributionell-statistische Information über das Auftreten bestimmter Laut- oder Silbenabfolgen.

Als segmentale Hinweise für die Strukturierung des Sprachstroms und die Identifizierung strukturell relevanter Einheiten kann zum einen akustische Information in Form von Assimilationsprozessen, die sich über eine Einheit – im wesentlichen die Silbe - erstrecken (Church, 1987), oder in Form von allophonischer Variation, die an bestimmte Positionen gebunden ist (Brent & Cartwright, 1996; Church, 1987; Nakatani & Dukes, 1977), dienen. Darüber hinaus kann sequentielle Information, d.h. Information über mögliche Abfolgen von Segmenten, Hinweise auf die interne Struktur einer sprachlichen Äußerung liefern. In diesem Zusammenhang wurden Übergangswahrscheinlichkeiten zwischen einzelnen Lautsegmenten (Cairns, Shillcock, Chater & Levy; 1997; Hayes & Clark, 1970) und zwischen Silben (Cowan, 1991; Saffran, Newport & Aslin, 1996) diskutiert. Weiterhin wurden phonotaktische Beschränkungen (Church, 1987) und im Sprachstrom häufig auftretende Phonemsequenzen bzw. Silben (Cowan, 1991; Morgan, Meier & Newport, 1987; Peters, 1985; Valian & Coulson, 1988; Valian & Levitt, 1996) als Strukturierungshilfe vorgeschlagen. Darüber hinaus kann auch bereits vorhandenes lexikalisches und syntaktisches Wissen als wissensgetriebener Strukturierungsprozess zur Analyse des sprachlichen Inputs beitragen (Christophe et al., 1997; Cole & Jakimik, 1980).

Grundsätzlich unterscheiden sich alle angesprochenen Informationstypen darin, ob ein bestimmter Hinweis auf das Vorliegen einer strukturellen Grenze direkt an der Grenze dieser potentiellen Einheit auftritt und somit diese direkt markiert oder ob sich

eine bestimmte phonologische Eigenschaft auf die strukturell zusammengehörenden Elemente verteilt, die dann durch diese gemeinsame Eigenschaft als Einheit markiert werden. Ersteres gilt beispielsweise für bestimmte allophonische Varianten. So sind im Englischen aspirierte Verschlusslaute auf die silbeninitiale Position beschränkt (Church, 1987), sie signalisieren also direkt eine Silbengrenze. Assimilationsprozesse signalisieren dagegen eine Einheit, da eine bestimmte akustische Eigenschaft durch die Assimilation auf eine größere Domäne projiziert wird.

Prinzipiell leisten natürlich beide Arten von Merkmalen das gleiche, da die Elemente zwischen direkten Grenzmarkierungen als Einheit aufzufassen sind und im Falle der Einheitenmarkierung, die Grenze der Einheit dort ist, wo die verbindende Eigenschaft nicht mehr auftaucht. Allerdings sind möglicherweise die Verarbeitung und damit eventuell auch das Lernen dieser verschiedenen Markierungsformen unterschiedlich. Es ist denkbar, dass lokale Grenzmarkierungen geringere Verarbeitungskapazitäten erfordern als Einheitenmarkierungen, die über größere Bereiche verteilt sind, da deren Ausnutzung eine parallele Verarbeitung der gesamten Einheit voraussetzt. Letzterer Punkt mag für silbeninterne Assimilationsprozesse noch nicht als problematisch einleuchten. Es liegen allerdings Hinweise darauf vor, dass erwachsene Sprecher gemeinsame Flexionsendungen einer Sequenz von Wörtern als Hinweis auf deren syntaktische Zusammengehörigkeit interpretieren (Morgan et al, 1987). Dieses Beispiel macht deutlich, dass die Domänen, deren Zusammengehörigkeit durch ein bestimmtes Merkmal signalisiert wird, relativ groß und über mehrere Wörter verteilt sein können. Für den erwachsenen Sprachverarbeiter mag es weniger bedeutsam sein, ob ein strukturelles Merkmal sich innerhalb einer Silbe befindet oder über größere Domänen innerhalb eines Satzes verteilt ist, bei Kinder ist die Größe der Domäne möglicherweise jedoch eine relevante Beschränkung: da ihre Verarbeitungskapazitäten wahrscheinlich noch nicht denen Erwachsener Sprecher entsprechen und offensichtlich zunächst an kleinere Domänen gebunden sind (Santelmann & Jusczyk, 1998), sind ihnen lokale Grenzmarkierungen wahrscheinlich zunächst leichter zugänglich.

Bevor auf eine mögliche Rolle dieser Art von Information im Erstspracherwerb eingegangen wird, soll in den folgenden Abschnitten wiederum zunächst geklärt werden, ob im sprachlichen Signal überhaupt Korrelationen der oben beschriebenen Art mit strukturellen Einheiten bzw. Grenzen existieren, ob diese prinzipiell von einem

Lernmechanismus entdeckt und für die Segmentierung bzw. den Erwerb struktureller Regularitäten ausgenutzt werden können und ob Sprecher/Hörer beim Verarbeiten und Lernen von Sprache diese Korrelationen nutzen. Dabei sind sowohl Befunde der deskriptiven Linguistik in Bezug auf die Sprachstruktur, als auch Ergebnisse von Modellierungen der Entdeckung und Nutzung dieser Hinweise anhand neuronaler Netze sowie psycholinguistischer Untersuchungen zur menschlichen Fähigkeit, diese Korrelationen zu entdecken und zu nutzen, von Bedeutung.

5.1.1 Akustische Hinweise

Nakatani und Dukes (1977) untersuchten die akustisch-phonetischen Eigenschaften von Phonemsequenzen, die sich lediglich darin unterschieden, an welcher Stelle die Grenze zwischen zwei Wörtern innerhalb dieser Sequenz war (z.B. *no notion* vs. *known ocean*). Wie ein Wahrnehmungsexperiment, in dem den Versuchspersonen diese Sequenzen präsentiert wurden und sie entscheiden sollten, welche Alternative sie gehört hatten, zeigte, variiert die Unterscheidbarkeit der Sequenzen in Abhängigkeit von der Qualität der Lautsegmente im Grenzbereich erheblich. Die von den Probanden gut unterscheidbaren Alternativen enthielten direkte Grenzmarkierungen in Form von allophonischen Varianten bzw. in Form von spezifischen Segmenten, die ausschließlich an den Silbengrenzen vorkommen. Beispielsweise ging wortinitialen Vokalen immer ein glottaler Verschlusslaut voran, den die Probanden auch klar als Merkmal für das Vorliegen einer Wortgrenze interpretierten. In den gut unterscheidbaren Sequenzen fanden sich zudem wortinitial häufig Verschlusslaute, die wesentlich stärker aspiriert als wortfinale Verschlusslaute. Bei den Konsonanten /r/ und /l/ waren deutliche allophonische Unterschiede zu erkennen, in Abhängigkeit davon, ob diese wortinitial oder wortfinal auftraten. Generell lässt sich aus den Ergebnissen dieser phonetischen Analysen schließen, dass sich segmentale Grenzmarkierungen im wesentlichen im initialen oder finalen Segment der kritischen Wörter finden und nicht über mehrere Silben verteilt sind.

Sehr deutliche Evidenz, dass segmentale Hinweise in Form von Allophonie zur Segmentierung ausgenutzt werden können, stammt aus Untersuchungen zur Vokalharmonie im Finnischen (Suomi, McQueen & Cutler, 1997; Vroomen,

Tuomainen & de Gelder, 1998). Das Finnische verfügt über drei Harmonieklassen, nämlich hintere Vokale /u a o/, vordere Vokale /y ? ? / und neutrale Vokale /i e/. Innerhalb eines Wortes dürfen keine hinteren und vorderen Vokale gemeinsam auftreten, sondern entweder nur Vokale einer Klasse oder Vokale aus einer dieser beiden Klassen gemeinsam mit einem neutralen Vokal. Das gilt sowohl für alle Stammvokale als auch für Suffixe, wodurch sich Fälle von Allomorphie ergeben. Die Vokalharmonie kann Hinweise auf Wortgrenzen geben, da ein Bruch in der Vokalharmonie bedeutet, dass zwischen den beiden Silben mit Vokalen aus verschiedenen Harmonieklassen eine Wortgrenze vorhanden sein muss. Tatsächlich können Sprecher des Finnischen diese Information auch ausnutzen. Sie können ein zweisilbiges Wort in einer dreisilbigen Sequenz schneller erkennen, wenn die nicht zum Wort gehörende Silbe gegenüber den Wortsilben die Vokalharmonie bricht, als wenn die dritte Silbe ebenfalls der Vokalharmonie entspricht (Suomi et al., 1997).

Sprecher/Hörer des Niederländischen und des Französischen – beides Sprachen, die keine Vokalharmonie aufweisen – zeigten dagegen beim selben Untersuchungsmaterial keine Nutzung der Vokalharmonie für Segmentierungsprozesse (Vroomen et al., 1998). Dieser Befund zeigt, dass der Effekt der Vokalharmonie nicht rein auf akustischen Merkmalen des Inputs basiert, sondern in Abhängigkeit von der sprachlichen Erfahrung des Hörers operiert. Es ist anhand vorliegender Befunde nicht zu entscheiden, ob finnische Sprecher aufgrund ihrer sprachlichen Erfahrung gelernt haben, dieses Merkmal zur Segmentierung zu nutzen, oder ob Sprecher des Niederländischen und des Französischen es verlernt haben, dieses Merkmal zu nutzen, da es in ihrer jeweiligen Muttersprache nicht von Bedeutung ist. Auch für die zuvor diskutierten akustischen Hinweise muss einschränkend bemerkt werden, dass sie wahrscheinlich nicht für eine reine datengetriebene Verarbeitung zu Beginn des Spracherwerbs geeignet sind, da beispielsweise positionale Beschränkungen für Allophone erst gelernt werden müssen.

5.1.2 Übergangswahrscheinlichkeiten zwischen Phonemen und Silben

Übergangswahrscheinlichkeiten zwischen Phonemen und Silben könnten eine wesentliche Rolle für die Erkennung wortähnlicher Einheiten im sprachlichen Input

darstellen. Wortformen sind gerade dadurch charakterisiert, dass sie eine fixe Abfolge von Phonemen und Silben darstellen, die konstant bleibt, unabhängig davon, in welchem Kontext sie auftreten. Innerhalb eines größeren Sprachkorpus, in dem dieselben Wortformen mehrfach auftauchen, lässt dies erwarten, dass die Übergangswahrscheinlichkeiten zwischen den Phonemen bzw. den Silben innerhalb der Wörter größer sind als zwischen Wörtern. Anders herum gesagt, Phoneme oder Silben, denen eine große Anzahl verschiedener Segmente folgen können, sind potentielle Endpunkte eines Wortes. Wenn dagegen nur eine relativ geringe Anzahl verschiedener Phoneme bzw. Silben einem gegebenen Segment folgen kann, ist die Wahrscheinlichkeit, dass diese zum gleichen Wort gehören relativ hoch. Diese Idee einer rein distributionellen Analyse geht zurück auf Harris (1955), der diese Vorgehensweise vorschlug, um ohne einen Zugriff auf semantische Informationen eine Morphemanalyse vornehmen zu können.

Inwieweit ein solcher Mechanismus, der rein auf der Berechnung von Übergangswahrscheinlichkeiten zwischen Segmenten innerhalb sprachlicher Äußerungen basiert, überhaupt in der Lage sein kann, Regelmäßigkeiten zu erkennen und diese mit Erfolg für die Erkennung von Wortgrenzen auszunutzen, wurde zunächst in erster Linie durch Computersimulationen überprüft. So stellten Cairns, Shillcock, Chater und Levy (1997) in einem phonetisch transkribierten Korpus englischer Sprache, in dem keine Wortgrenzen markiert waren, fest, dass ein Segmentierungsmechanismus, der anhand der Analyse der im Korpus selbst auftretenden Konsonantenfolgen potentielle Segmentierungspunkte sucht, überzufällig häufig Wortgrenzen korrekt identifizieren kann. Dieser Mechanismus berechnete zunächst die Häufigkeit aller im Korpus vorkommenden Folgen von zwei bzw. drei Phonemen. Diese Häufigkeiten wurden anschließend verwendet, um für jedes Phonem in einem Testkorpus das folgende vorherzusagen. Falsche Vorhersagen wurden als potentielle Wortgrenzen angesehen. Ohne jede weitere Zusatzinformation war das System in der Lage, ca. 20% der Wortgrenzen des Korpus korrekt zu ermitteln, wobei das Verhältnis von korrekter Zuweisung einer Wortgrenze (Treffer) zur falschen Zuweisung einer Wortgrenze (falscher Alarm) 1.5 zu 1 lag. Die Erfolgsquote erhöhte sich, wenn, wenn Pausen und Sprecherwechsel aus dem ursprünglichen Korpus als weitere Informationsquellen in den Lernkorpus eingefügt wurden: in diesem Fall lag die korrekte Identifikation von

Wortgrenzen im Testkorpus bei ca. 30% mit einem Verhältnis von Treffern und falschem Alarm von 2.4 zu 1. Die Verbesserung der Trefferquote unter diesen Bedingungen ist natürlich darauf zurückzuführen, dass Pausen und Sprecherwechsel im allgemeinen nicht innerhalb von Wörtern stattfinden, sondern an Wortgrenzen.

Zu einem ähnlichen Ergebnis kommen Christiansen, Allen und Seidenberg (1998), die Segmentierungsprozesse anhand der Ausnutzung verschiedener Hinweise an einem realen Korpus kindgerichteter Sprache simulierten. Dabei wurden verschiedene Trainingskorpora verwendet, die neben einer phonemischen Transkription für die Berechnung von Übergangswahrscheinlichkeiten zwischen Phonemen auch Information über Äußerungsgrenzen enthalten konnten. Im Testkorpus sollte das System für jedes Phonem das wahrscheinlich folgende ermitteln, wobei auch Wortgrenzen vorausgesagt werden konnten. In der Bedingung, in dem der Trainingskorpus neben der phonemischen Transkription auch Information über die Äußerungsgrenzen enthielt, waren 37% aller vorhergesagten Wortgrenzen korrekt und 40% der Wörter des Korpus waren korrekt segmentiert, d. h. sowohl die finale als auch die initiale Wortgrenze waren korrekt, ohne dass innerhalb des Wortes eine Genze zugewiesen worden war. In dieser Analyse zeigt sich außerdem, dass ein auf der Häufigkeit von Phonemfolgen basierender Segmentierungsmechanismus solche Wörter besonders zuverlässig korrekt zu segmentieren lernt, die im Input eine hohe Gebrauchshäufigkeit aufweisen. So war im Trainingskorpus das Pronomen *you* das mit Abstand häufigste Wort, gerade dieses wurde im Testkorpus mit hoher Korrektheit segmentiert.

Die Bedeutsamkeit von Äußerungsgrenzen als zusätzlicher Informationsquelle für das Erlernen phonotaktischer Hinweise auf Wortgrenzen ergibt sich wiederum daraus, dass eine Äußerungsgrenze auch immer eine Wortgrenze darstellt, so dass die Phoneme oder Phonemkombinationen, die in unmittelbarer Nachbarschaft einer Äußerungsgrenze auftreten, immer auch potentielle Wortanfänge bzw. Wortenden sind. Die Markierung von Äußerungsgrenzen bzw. Satzgrenzen spiegelt auch durchaus die realen Verhältnisse im akustischen Input für den Hörer wieder, da Äußerungsgrenzen im allgemeinen mit Grenzen von Intonationsphrasen zusammenfallen und daher akustisch gut markiert sind.

Diese Modellierungen zeigen übereinstimmend, dass natürlich-sprachliche Äußerungen distributionelle Informationen enthalten, die zur Segmentierung der

Äußerungen in Wörter beitragen kann. Diese Segmentierungsmechanismen sind auch relativ erfolgreich, je nach verwendetem Algorithmus und der Menge der verschiedenen Informationen, die dieser berücksichtigt, werden weitaus mehr korrekte Identifikationen von Wortgrenzen berichtet, als dies nach einer rein zufälligen Zuweisung von Wortgrenzen der Fall wäre. Aus der Tatsache, dass natürliche Sprache Informationen dieser Art enthält und diese prinzipiell auch zu einer erfolgreichen Segmentierung beitragen können, lässt sich jedoch nicht zwangsläufig schließen, dass das menschliche Verarbeitungssystem diese Information auch ausnutzen kann bzw. tatsächlich ausnutzt. Es ist fraglich, ob die Ergebnisse der Computersimulationen auf die Situation der natürlichen Sprachverarbeitung übertragbar sind. So sind die Trainingskorpora, die in den Modellierungen benutzt wurden, relativ groß. Beispielsweise umfasst der Trainingskorpus kindgerichteter Sprache, der von Christiansen et al. (1998) verwendet wurde, ca. 25.000 Token verschiedener Wortformen in ca. 8.000 Äußerungen. Für die Bestimmung der distributionellen Merkmale dieses Korpus muss das System die Vorkommenshäufigkeit jeder einzelnen in dem Korpus vorkommenden Phonemabfolge berechnen.

Es ist kaum davon auszugehen, dass diese Kapazitäten dem menschlichen Verarbeitungssystem vergleichbar sind, d. h. dass es in der Lage ist, diese großen Informationsmengen parallel zu verarbeiten bzw. über einen längeren Zeitraum im Gedächtnis zu behalten. Dies ist jedoch eine der zentralen Voraussetzungen für eine distributionelle Analyse wie sie simuliert wurde. Grundsätzlich ist das menschliche Verarbeitungssystem allerdings in der Lage, distributionelle Eigenschaften des Inputs zu berechnen: dies wurde bislang allerdings nur bei der konzentrierten Präsentation von nur wenigen Stimuli mit klaren Distributionsmustern nachgewiesen.

Im Gegensatz zu den bisher dargestellten Computersimulationen wurden in Untersuchungen zur menschlichen Fähigkeit, Kookkurrenzmuster in sprachlichen Reizen zu erkennen, bislang ausschließlich Silben als grundlegende Einheit verwendet. In ihrer Methodik basieren diese Arbeiten auf einer Untersuchung von Hayes und Clark (1970), die die menschliche Fähigkeit, distributionelle Regelmäßigkeiten zu erkennen, mit nonverbalen akustischen Reizen in Form von Tonsequenzen überprüften. Sie präsentierten ihren Versuchspersonen zunächst einen 45-minütigen kontinuierlichen Lautstrom, der vier verschiedene Tonsequenzen in variierenden Abfolgen enthielt. Im

Anschluss daran wurden den Probanden Teile des Kontinuums präsentiert, in die entweder Pausen an den Grenzen der Vierersequenzen oder innerhalb der Vierersequenzen eingefügt worden waren. Es zeigte sich, dass die Probanden, die Tonfolgen, die Pausen an den Grenzen der Vierersequenzen aufwiesen, den zuerst gehörten Sequenzen als ähnlicher einschätzten, als die Tonfolgen, in denen die Pausen innerhalb der Vierersequenzen auftraten.

Cowan (1991) ging der Frage nach der Rolle der Häufigkeit der Kookkurrenzen von Silben bei der Strukturierung von Silbenketten nach. In seiner Untersuchung wurden den Probanden zunächst zwei verschiedene Silbensequenzen präsentiert, die entweder in einer Silbe übereinstimmten (*wo-wie*, *wie-bi-gra*) oder in zwei Silben (*ga-li-se*, *li-se-stu*). Variiert wurde die Häufigkeit, mit der die beiden Silbensequenzen präsentiert wurden: entweder wurden beide gleich häufig präsentiert oder eine der Silbensequenzen wurde doppelt so häufig präsentiert wie die andere. Die Aufgabe der Versuchspersonen bestand darin, in einer im Anschluß präsentierten Vierersequenz (*ga-li-se-stu*, *wo-wie-bi-gra*) alle strukturellen Grenzen zu markieren. Zumindest bei den Sequenzen, die lediglich in einer einzigen Silbe übereinstimmten zeigte sich ein deutlicher Effekt der Präsentationsfrequenz: die Probanden betrachteten die Subsequenzen, die ihnen zuvor häufiger präsentiert worden waren, als stärkere strukturelle Einheit als die seltener präsentierten Subsequenzen. Bei den Dreisilblern tendierten die Probanden dazu, die zweisilbige Subsequenz, die in allen Präsentationen vorgekommen war (*li-se*) als strukturelle Einheit zu markieren. Hier spielte die Variation der Häufigkeit der Präsentation der Gesamtsequenz keine Rolle.

Dieser Befund ist deshalb bemerkenswert, da er darauf hindeutet, dass die Strategie der Probanden eher darauf ausgerichtet war, zusammengehörende Einheiten in diesen Silbensequenzen zu suchen, als einzelne Silben herauszuextrahieren. Im zweiten Fall wäre nämlich bei der Zuweisung von Grenzen in die Vierersequenz zu erwarten gewesen, dass der einzelnen Silbe, die in allen vorher präsentierten Segmenten vorkam, rechts und links Grenzmarkierungen zugewiesen worden wären. Dieser Befund könnte ein Hinweis darauf sein, dass das menschliche Verarbeitungssystem eher einheitensuchend als grenzsuchend operiert. Dieser Schluss bedarf jedoch weiterer Überprüfung, da auch andere Interpretationen dieses Musters möglich sind. So könnte man auch vermuten, dass eine einzelne CV-Silbe kein typisches Wort darstellt und

daher die Zweisilbler als Einheiten heraussegmentiert wurden, nicht aber die einzelne Silbe. Tatsächlich entspricht eine einzelne CV-Silbe mit kurzem Vokal nicht den Beschränkungen für den minimalen phonologischen Aufbau eines Wortes im Englischen. Auch hier zeigt sich wieder, dass in Strukturierungsentscheidungen verschiedene Typen von Informationen einbezogen werden.

Saffran, Newport und Aslin (1996) verfolgten diesen Strang der Nutzung von Kookkurrenzmustern von Silben bei erwachsenen Sprechern/Hörern weiter. Sie plädieren allerdings dafür, dass nicht nur die Häufigkeit des Aufeinanderfolgens zweier Elemente relevant ist, sondern dass die Stärke der Abfolgeinformation als Hinweis auf eine strukturelle Einheit daran gemessen werden muss, wie häufig das erste Element der Abfolge in einem anderen Kontext vorkommt, d. h. in einer Sequenz mit einem anderen Laut bzw. einer anderen Silbe. Diese - an der absoluten Häufigkeit des ersten Elements relativierte - Häufigkeit des Aufeinanderfolgens zweier Elemente bezeichnen die Autoren als Übergangswahrscheinlichkeit. Für ihre Untersuchung konstruierten Saffran und Mitarbeiter sechs dreisilbige Pseudowörter, die ausschließlich aus CV-Silben bestanden. Ein Teil der CV-Silben wurde in mehreren der Pseudowörter verwendet, so dass die Übergangswahrscheinlichkeiten der Silben innerhalb der Wörter unterschiedlich hoch war (zwischen 30% und 100%). Diese Pseudowörter wurden zu kontinuierlichen Sprachstrom zusammengefügt, in dem die dreisilbigen Pseudowörter in zufälligen, wechselnden Abfolgen wiederholt wurden. Dieses Sprachkontinuum wurden den Untersuchungsteilnehmern zu Beginn des Experiments präsentiert. Für die eigentliche Testphase wurden neue Silbensequenzen konstruiert, von denen die Hälfte aus Kombinationen bestand, die im kontinuierlichen Sprachstrom der Eingangsphase niemals direkt hintereinander aufgetreten waren und die andere Hälfte aus einem Silbenpaar aus den Pseudowörtern der Eingangsphase und einer weiteren Silbe. Diese neuen Silbensequenzen wurden den Probanden in der Testphase paarweise zusammen mit den dreisilbigen Pseudowörtern aus dem Sprachstrom präsentiert, wobei die Probanden jeweils entscheiden sollten, welche Sequenz einem Wort der zuvor gehörten Sprachsequenz entsprach.

Es zeigte sich, dass die Leistungen oberhalb des Rateniveaus lagen, d. h. die Probanden hatten die Wörter aus dem kontinuierlichen Sprachstrom gelernt, wobei die Leistungen bei den Paaren, in denen die noch nicht gehörte Silbensequenz ein

Silbenpaar der eigentlichen Wörter enthielt etwas schlechter waren als bei den Paaren, in denen die neue Sequenz ausschließlich noch nicht gehörte Silbenkombinationen enthielt. Es zeigte sich zudem ein Effekt der Variation der Übergangswahrscheinlichkeit innerhalb der Wörter: die Probanden erkannten die Wörter mit hoher interner Übergangswahrscheinlichkeit der Silben mit höherer Sicherheit als die Wörter mit niedrigerer interner Übergangswahrscheinlichkeit der Silben.

Diese Befunde zeigen übereinstimmend, dass das menschliche Verarbeitungssystem durchaus in der Lage ist, zumindest auf Silbenebene die statistischen Gegebenheiten eines längeren sprachlichen Reizes zu berechnen. Diese Information wird offensichtlich auch herangezogen, um kontinuierliche Signale zu segmentieren. Es ist aber noch einmal zu betonen, dass die Menge der verwendeten Reize in den Untersuchungen sehr gering war und diese Reize in einer längeren Präsentation sehr häufig wiederholt wurden. Ob sich diese Mechanismen deshalb ebenso gut für die Segmentierung natürlicher Sprache eignen, muss deshalb zunächst offen bleiben.

5.1.3 Phonotaktische Beschränkungen

Phonotaktische Beschränkungen im klassisch linguistischen Sinne hängen sehr eng mit den distributionellen Übergangswahrscheinlichkeiten zwischen Phonemen zusammen, da sie Phonemsequenzen ausschließen, die in einer Sprache überhaupt nicht oder in einer bestimmten Silben- oder Wortposition nicht auftreten. In dieser Hinsicht sind phonotaktische Beschränkungen strukturelle Bedingungen, die bestimmte Abfolgemoöglichkeiten von Phonemen regeln, und nicht wie die vorher behandelten Übergangswahrscheinlichkeiten rein statistische Eigenschaften des Sprachgebrauchs. Der enge Zusammenhang zwischen beiden besteht jedoch wahrscheinlich im Erwerb phonotaktischer Regeln. Da diese hochgradig einzelsprachspezifisch sind, muss der Lerner sie aus dem sprachlichen Input herausfinden. Als phonotaktisch nicht zulässig kann der Lerner danach solche Kombinationen klassifizieren, die er in seinem Input nicht vorfindet. Da phonotaktische Beschränkungen nur innerhalb bestimmter Domänen gelten, im allgemeinen die Silbe, kann phonotaktische Information für die Bestimmung der Grenzen dieser Domänen genutzt werden, d. h. findet sich im Input eine

phonotaktisch nicht zulässige Phonemfolge, so kann davon ausgegangen werden, dass sich innerhalb dieser Folge eine strukturelle Grenze befindet.

Befunde von McQueen (1998) weisen darauf hin, dass erwachsene Hörer phonotaktische Beschränkungen tatsächlich für die Wortsegmentierung ausnutzen können. Er stellte fest, dass Sprecher des Niederländischen in zweisilbige Pseudowörter eingebettete einsilbige Wörter dann schneller erkannten, wenn das Wortende bzw. der Wortanfang mit einer Silbengrenze korrelierte, die aus phonotaktischen Gründen an dieser Stelle obligatorisch war. So konnten die Probanden beispielsweise das Wort *pil* schneller erkennen, wenn es in die Sequenz *pilrem* eingettet war, in der phonotaktisch bedingt zwischen dem /l/ und dem /r/ eine Silbengrenze sein muss. Demgegenüber ergaben sich in einer Sequenz wie *pilvrem*, in der die Silbengrenze nach phonotaktischen Beschränkungen sowohl nach dem /l/ als auch nach dem /v/ auftreten könnte, verzögerte Latenzen.

Eine Schwäche der Ausnutzung phonotaktischer Information für die Bestimmung von Wortgrenzen könnte darin bestehen, dass – wie schon gesagt – die eigentliche Domäne phonotaktischer Beschränkungen die Silbe und nicht das gesamte Wort ist, d. h. zuverlässige Information über Wortgrenzen kann phonotaktische Information lediglich für einsilbige Wörter liefern. In Sprachen wie dem Englischen mit einem relativ hohen Anteil einsilbiger Wörter – so fanden beispielsweise Christiansen et al. (1998) in ihrem kindgerichteten Korpus, dass 86% aller Worttoken einsilbig waren – könnte diese Information über Silbengrenzen jedoch für die Wortsegmentation sehr hilfreich sein. In stärker flektierenden Sprachen, wie beispielsweise dem Deutschen, ist dagegen sicherlich der Anteil einsilbiger Wörter geringer, so dass phonotaktische Information zur Identifizierung von Wortgrenzen nicht sehr effizient beitragen kann.

In den Bereich der phonotaktischen Beschränkungen gehören auch generelle Bedingungen über den minimalen phonologischen Aufbau einer Wortform, sogenannte *minimal word constraints*. Beispielsweise geht in den von Brent und Cartwright (1996) vorgeschlagenen Segmentierungsalgorithmus Wissen dieser Art ein, nämlich beispielsweise die Beschränkung, dass in jedem Wort mindestens ein Vokal vorhanden sein muss. Diese Beschränkung verwirft beispielsweise solche Segmentierungshypothesen, in denen eine Kette von Konsonanten als eigenständige

Segmentfolge und damit als Wortform angenommen wird (vgl. Norris, McQueen, Cutler & Butterfield, 1997).

Befunde von Norris und Mitarbeitern (1997) weisen daraufhin, dass diese Beschränkung tatsächlich für die Sprachverarbeitung relevant ist. Auch in ihrem Experiment hatten die Versuchspersonen die Aufgabe, ein in ein Pseudowort eingebettetes Wort zu identifizieren. In der Untersuchung gab es zwei Bedingungen. In der ersten konnte der Kontext des eingebetteten Wortes nach phonotaktischen Gesichtspunkten selbst ein Wort darstellen (z.B. *vuffapple*). In der zweiten Bedingung blieb nach der Segmentierung des Wortes nur noch ein einzelner Konsonant übrig, der nach phonotaktischen Restriktionen kein eigenständiges Wort darstellen kann (z.B. *fapple*). In den Ergebnissen zeigten sich schnellere Reaktionen und weniger Fehler in der ersten Bedingung, d. h. offensichtlich war es einfacher, das integrierte Wort zu entdecken, wenn die übrigbleibenden Segmente ihrerseits selber zu einer potentiellen Wortform integriert werden konnten. Befunde mit Sprechern des Japanischen zeigen, dass der gefundene Unterschied nicht nur mit der Menge der Restsegmente zu erklären ist, sondern tatsächlich auf dem möglichen Wortstatus der Restmenge beruht. Die japanischen Sprecher/Hörer zeigten bei gleicher Aufgabenstellung deutliche Unterschiede zwischen Stimuli, in denen der überzählige Laut ein Konsonant war, und Stimuli, in denen dieser ein Vokal war (McQueen, Otake & Cutler, 2001). Dabei fiel es den Probanden leichter, das eingebettete Wort zu entdecken, wenn der Kontextlaut ein Vokal war. Da im Japanischen auch einzelne Vokale ein Wort darstellen können, stützt auch dieser Befund die Rolle eines *minimal-word-constraints* in der Sprachverarbeitung.

5.1.4 Häufige Einheiten

In diesem Abschnitt geht es um die Rolle, die im Input häufig auftretende einzelne Sequenzen von Phonemen oder Silben für die Strukturierung des Inputs haben können. Im Gegensatz zu Abschnitt 5.1.2 geht es hierbei nicht um die Frage, ob der Hörer sensitiv gegenüber Übergangswahrscheinlichkeiten zwischen einzelnen Elementen seines sprachlichen Inputs ist und diese Übergangswahrscheinlichkeit zur Bildung von Einheiten nutzt, sondern es geht um die Frage, inwieweit Einheiten, die bereits als

solche identifiziert wurden, zur weiteren Strukturierung des Inputs herangezogen werden können. Von der im nächsten Abschnitt folgenden Betrachtung lexikalischer Mechanismen wird dies zunächst getrennt, da in diesem Zusammenhang noch nicht von Wortformen gesprochen werden soll, sondern lediglich von Phonem- oder Silbenmustern, die im Input häufig sind. Diese können natürlich auch hochfrequenten Wörtern entsprechen, es können aber auch einfache Phonemsequenzen, wie beispielsweise Affixe sein. Wie schon erwähnt, werden solche hochfrequenten Phonemfolgen bei Simulationen der Berechnung der Häufigkeit von Phonemsequenzen und der Nutzung dieser Information für weitere Segmentierungsprozesse sehr schnell als eigenständige Formen extrahiert und segmentiert (Christiansen et al., 1998). Die Nützlichkeit solcher wiederkehrender Muster wurde in erster Linie in Experimenten mit erwachsenen Sprechern zum Lernen von Kunstsprachen, vor allen Dingen von syntaktischen Aspekten der Kunstsprache, überprüft. Dies ist ein weiterer Grund, warum diese Thematik gesondert von tatsächlichen lexikalischen Mechanismen behandelt wird.

In einer ganzen Reihe von Arbeiten zum Erwerb künstlicher Grammatiken bei Erwachsenen wurde der Frage nachgegangen, inwieweit sogenannte *Marker*, d. h. formstabile Elemente, die in ihren phonologischen Eigenschaften Funktionswörtern und Flexionsendungen nachempfunden waren, also beispielsweise kürzer waren als die „Inhaltswörter“ der Kunstsprache, das Lernen syntaktischer Eigenschaften der Kunstsprache fördern.

Green (1979) untersuchte, ob *Marker*, die konsistent analog zu einem Funktionswort vor einem Inhaltswort dessen Klassenzugehörigkeit markieren, bzw. an einer Phrasengrenze den Phrasentyp markieren, für das Lernen der Syntax einer Kunstsprache genutzt werden. Nach seinen Befunden lernen Erwachsene die Regularitäten einer Kunstsprache, in der *Marker* in der genannten Weise systematisch eingesetzt sind, effizienter als eine Kunstsprache, in der keine *Marker* verwendet werden bzw. in denen sie unsystematisch auftauchen.

Gleiche Befunde berichten Mori und Moeser (1983) für das Lernen einer Kunstsprache, in der die *Marker* wie Suffixe den Inhaltswörtern folgten und gleiche Suffixe auf gleiche Wortklassenzugehörigkeit hinwies. Interessanterweise zeigte sich kein Einfluss dieser *Marker* mehr, wenn den Wörtern der Kunstsprache Referenten

zugewiesen worden waren und sich die syntaktischen Regeln aus diesem Referenzfeld ableiten ließen. In diesem Fall zeigten die Versuchspersonen keine Leistungsunterschiede zwischen einer Bedingung, in der Referenzfeld und *Marker* konsistent die syntaktischen Regeln markierten und einer Bedingung, in der Referenzfeld und *Marker* inkonsistent gebraucht wurden. Im zweiten Fall verliessen sich die Versuchspersonen konsistent auf das Referenzfeld.

Die Annahme, dass Erwachsene verschiedene Typen von Information ausnutzen können – je nachdem welche im Input vorhanden ist bzw. dominiert – wird auch von Befunden von Meier und Bower (1986) unterstützt. Sie kombinierten in einer Inputbedingung ihre Sprache – englische Wörter, deren Syntax jedoch in der Wortfolge vom Englischen abwich – nur mit einem Referenzfeld, d. h. einer bildlichen Darstellung der Inhalte des Satzes. In einer zweiten Bedingung fehlte das Referenzfeld, die Zusammengehörigkeit von Phrasen war jedoch durch die Markierung aller Wörter einer Phrase mit einem suffixähnlichen formgleichen *Marker* gekennzeichnet. Gegenüber einer Kontrollbedingung, in der beide Informationstypen fehlten, waren die Lerner dieser beiden Bedingungen gleich erfolgreich, d. h. die semantische Information stützte das Lernen in gleichem Ausmaß wie die morpho-syntaktische. Diese Befunde zeigen deutlich, dass das Verarbeitungs- und Lernsystem sich sehr flexibel den verschiedenen Inputbedingungen anpassen kann.

Die Nutzung verschiedener Typen von *Markern* verglichen Morgan, Meier und Newport (1987). Sie benutzten in ihrer Kunstsprache neben prosodischer Information zur Markierung von Phrasen sowohl freie *Marker*, die den Phrasenbeginn markierten, als auch *Suffix-Marker*, die in einheitlicher Form die in einer Phrase zusammengehörenden Elemente markierten. Auch unter Verwendung eines Referenzfeldes, zeigten sich bei allen strukturellen Hinweisen bessere Lerneffekte in Bezug auf die Syntax der Kunstsprache, als beim Fehlen bzw. bei einer inkonsistenten Verwendung der Hinweise.

Eine ähnliche Vorgehensweise findet sich in einer Untersuchung von Valian und Coulson (1988). Sie überprüften, welche Rolle die Variation der Frequenz einer Phonemfolge im Stimulusmaterial für das Lernen von Regularitäten hat. In ihrer Kunstsprache wurden den Probanden Sätze, die jeweils aus zwei Phrasen bestanden, präsentiert. Die Phrasen selbst bestanden aus je zwei Wörtern, von denen eins – von den

Autoren als *Marker* bezeichnet – nur aus drei Buchstaben bestand und initial einen Vokal aufwies. Das andere Wort, von den Autoren als Inhaltswort bezeichnet, war mit vier bis sechs Buchstaben länger und wies einen initialen Konsonanten auf. Der *Marker* ging dem Inhaltswort in den Phrasen stets voran. Es wurden zwei Phrasentypen unterschieden (A und B), die jeweils spezifische *Marker* bzw. Inhaltswörter enthielten, d.h. Phrasentyp A enthielt *Marker* und Inhaltswörter einer anderen Kategorie als Phrasentyp B. Als Trainingssätze wurden 24 Sätze aus jeweils zwei Phrasen erstellt. Dabei wurden zwei „Dialekte“ der Kunstsprache konstruiert, die sich in der Frequenz der Verwendung von Formen der verschiedenen Kategorien unterschieden. Im hochfrequenten Dialekt wurden innerhalb jeder *Marker*-Kategorie nur zwei verschiedene Formen benutzt und innerhalb jeder Inhaltswortkategorie sechs verschiedene Formen. Daraus resultiert, dass jeder einzelne *Marker* in den insgesamt 24 Trainingssätzen relativ häufig vorkommt, die einzelnen Inhaltswörter sehr viel seltener. Im niedrigfrequenten Dialekt wurden aus jeder *Marker*-Kategorie vier verschiedene Formen verwendet und aus jeder Inhaltswortkategorie drei. Damit ist kein deutlicher Unterschied in der Auftretenshäufigkeit zwischen einzelnen Elementen der Kunstsprache mehr gegeben.

Nach dem Lesen der Trainingssätze wurden den Probanden Testsätze präsentiert, die in ihrem Aufbau entweder den Trainingssätzen entsprachen oder Verstöße gegen die zugrundeliegenden Konstruktionsregeln, beispielsweise Wortfolgefehler, aufwiesen. Die Probanden hatten die Aufgabe, bei jedem Satz zu entscheiden, ob der Satz den Trainingssätzen ähnlich war oder nicht. Es zeigte sich, dass die Lerner des hochfrequenten Dialekts bei dieser Aufgabe weitaus weniger Fehler machten als die Lerner des niedrigfrequenten Dialekts. Dieser Befund zeigte sich auch in einer teilweisen Replikation der Studie bei auditiver Präsentation des Stimulusmaterials (Valian & Levitt, 1996).

Die Ergebnisse zeigen deutlich, dass die Auftretenshäufigkeit von Elementen einer Kategorie vom Lerner genutzt wird. Offensichtlich gelingt es den Lernern besser zu erkennen, dass in der Kunstsprache Elemente zweier Kategorien verwendet werden, die in bestimmten Reihenfolgen auftreten und in bestimmten Abhängigkeiten zueinander stehen, wenn die Menge der in beiden Kategorien verwendeten konkreten Formen sehr unterschiedlich ist, d. h. die Elemente einer Kategorie eine relativ hohe

Frequenz aufweisen gegenüber den Elementen der anderen Kategorie. Auf natürliche Sprachen übertragen, weist dieses Ergebnis einen Weg zumindest zu einer groben Kategorisierung zwischen Funktions- und Inhaltswörtern. Eines der charakteristischen Merkmale von Funktionswörtern besteht gerade in ihrer relativ hohen Frequenz gegenüber Inhaltswörtern (z.B. Morgan, Shi & Allopenna, 1996). Wie die Untersuchungen von Valian und Mitarbeitern zeigen, könnte dieses Merkmal auch zu einem relativ schnellen Aufbau lexikalischer Repräsentationen für diese Elemente führen.

5.1.5 Lexikalisches Wissen

In den meisten gängigen Modellen der auditiven Worterkennung wird das Problem der Wortsegmentierung nicht als eigenständiger Prozess innerhalb der Worterkennung behandelt, sondern schlichtweg als das Resultat der lexikalischen Verarbeitung, so beispielsweise im Kohorten-Modell (Marslen-Wilson, 1987, 1990). Nach diesem Modell ergibt sich die Erkennung von Wortgrenzen im sprachlichen Signal durch die Identifizierung einer lexikalisch-phonologischen Repräsentation, die mit dem Input kompatibel ist. Eine der wesentlichen Annahmen des Kohorten-Modells besteht darin, dass die Erkennung eines Wortes bereits abgeschlossen werden kann, ohne dass die gesamten Segmente des Wortes aus dem Input verarbeitet werden müssen. Dies wird geleistet durch das Zusammenwirken von datengetriebenen *bottom-up* Prozessen und wissensgetriebenen *top-down* Prozessen. Nach diesem Modell wird anhand der initialen akustischen Information aus einem Wort zunächst eine Kohorte von Wortformen im Lexikon aktiviert, deren phonologische Information mit der akustischen Information kompatibel ist. Fortlaufend wird diese initiale Kohorte anhand weiterer Inputinformation aber auch anhand von Kontextinformation reduziert, bis nur noch ein Kandidat der Ausgangskohorte übrigbleibt. Dies ist der Punkt der Worterkennung. Empirische Evidenz für die Aktivierung von Konkurrenten mit übereinstimmenden initialen Segmenten während der auditiven Worterkennung liefern beispielsweise die Experimente von Zwitserlood (1989). In Abhängigkeit vom phonologischen Aufbau des Wortes und dem Isolationspunkt im Wort, d.h. dem Punkt ab dem es keine weiteren phonologisch übereinstimmenden Konkurrenzwörter mehr im Lexikon gibt, und dem

Kontext, in dem ein Wort verwendet wird, kann der Erkennungspunkt für ein Wort also schon sehr weit am Anfang des Wortes liegen. Mit der Erkennung des Wortes wird - wie gesagt - dessen phonologische Repräsentation verfügbar und somit auch phonologische Information über das Wortende. Anhand dieser Information ist der Verarbeitungsmechanismus dann in der Lage, das Wortende im akustischen Stimulus zu identifizieren und somit auch den Beginn des folgenden Wortes. Nach diesem Modell ist also die Ausgangskohorte bestimmt durch die phonologische Übereinstimmung der initialen Segmente der lexikalischen Einträge, danach ist die Identifizierung des Wortanfangs eine wesentliche Voraussetzung für die lexikalischen Verarbeitungsprozesse.

Auch im sogenannten TRACE-Modell (McClelland & Elman, 1986) wird auf einen expliziten prälexikalischen Segmentierungsprozess verzichtet. Dieses Modell unterscheidet sich vom Kohorten-Modell darin, dass wortinitialen Segmenten kein besonderer Status in den lexikalischen Verarbeitungsprozessen eingeräumt wird. So werden alle lexikalischen Einträge, die phonologisch mit dem Input-Signal übereinstimmen während der lexikalischen Prozesse aktiv. Hierbei ist es nicht relevant, in welcher Wortposition diese phonologische Übereinstimmung auftritt. Die Segmentierung des Inputs ergibt sich dann erst daraus, dass die lexikalischen Kandidaten ausgewählt werden, die den höchsten Grad an Aktivierung erreichen (vgl. Frauenfelder & Peeters, 1990). Hier können zusätzlich zur Input-Information bestimmte Beschränkungen über mögliche Wortformen und generelle Verarbeitungsprinzipien, beispielsweise, dass alle Segmente des Inputs einer Wortform zugewiesen werden müssen, zum tragen kommen (vgl. Brent & Cartwright, 1996).

Dass solche lexikalischen *top-down* Prozesse tatsächlich eine Rolle spielen, ist zum einen an Frequenzeffekten in Monitoring-Aufgaben zu beobachten. So hat man beispielsweise gefunden, dass in Phonem-Monitoring-Aufgaben, die Schnelligkeit mit der ein Phonem in einem Wort erkannt werden kann, unter anderem mit der Gebrauchshäufigkeit des Wortes variiert, das dem targetenthaltenden Wort vorangeht (Foss & Blank, 1980). Außerdem weiß man aus einer Vielzahl von Untersuchungen, dass hochfrequente Wörter schneller erkannt werden als Wörter mit niedrigerer Gebrauchshäufigkeit (z.B. Marslen-Wilson, 1990; Taft & Hambly, 1986). Danach könnte dieser Effekt wie folgt interpretiert werden: Ein Wort mit hoher Frequenz wird

schnell erkannt, demnach steht schnell dessen phonologische Repräsentation zur Verfügung und somit Information über den Beginn des folgenden Wortes, d. h. die lexikalischen Verarbeitungsprozesse für das folgende Wörter können schneller initiiert werden.

Direkte Evidenz für diese Annahme findet sich in einer Studie von Christophe et al. (1997). Die Forscher führten mit erwachsenen Sprecher/Hörern zwei Varianten einer Phonem-Monitoring Aufgabe durch, bei der die Probanden immer dann, wenn in einem akustischen Stimulus ein vorher spezifiziertes Phonem auftrat, auf eine Reaktionstaste drücken sollten. In der einen Variante der Aufgabe sollten die Versuchspersonen bei jedem Auftreten des Zielphonems reagieren, in der zweiten Variante nur dann, wenn das Zielphonem wortinitial auftrat. Die Stimuli bestanden aus Nominalphrasen, bestehend aus einem Determinierer, einem Nomen und einem Adjektiv, wobei das Adjektiv stets dem Nomen folgte. Die initialen Zielphoneme waren entweder das erste Phonem des Nomens oder das erste Phonem des Adjektivs. Die Annahme der Forscher war, dass die Identifikation des Wortanfangs des Nomens durch den vorangehenden Determinierer einfacher sein sollte, als die Identifikation des Anfangs des Adjektivs und sich dies vor allen Dingen in Unterschieden in den Reaktionszeiten in der Aufgabe auswirken sollte, in der die Probanden ausschließlich auf die wortinitialen Zielphoneme reagieren sollten. Tatsächlich deutet sich in den Ergebnissen dieser Unterschied auch an. Während sich bei den Nomen kein Unterschied in den Reaktionszeiten zwischen den beiden Aufgabentypen fand, waren bei den Adjektiven die Reaktionszeiten in der positionsunabhängigen Aufgabe signifikant kürzer als die Reaktionszeiten in der positionsbeschränkten Aufgabe. Die Ursache hierfür sehen die Autoren darin, dass es im Fall der Adjektive schwieriger war den Wortanfang innerhalb der Phonemsequenz zu erkennen als im Falle der Nomen. Leider war in dem Experiment die Wortklasse mit dem Vorangehen eines Funktionswortes konfundiert, da den Nomen immer ein Funktionswort voranging und den Adjektiven nie, so dass nicht ganz ausgeschlossen werden kann, dass der gefundene Unterschied ein Wortklasseneffekt ist.

Für die Satzverarbeitung scheinen Funktionswörter mit ihrer syntaktisch-lexikalischen Information darüber hinaus als Markierungen der syntaktischen Struktur einer Äußerung von Bedeutung zu sein. So zeigt beispielsweise eine Vielzahl von Studien, dass im Englischen mehrfach eingebettete Relativsätze ohne Relativpronomen

schwerer zu verarbeiten sind als Sätze mit Relativpronomen (Fodor & Garrett, 1967; Hakes & Cairns, 1970; Hakes, Evans & Brannon, 1976) und dass Komplementsätze mit einleitendem Komplementierer einfacher zu verarbeiten sind als nicht-eingeleitete Komplementssätze (Hakes, 1972). Dieser Effekt ist wahrscheinlich darauf zurückzuführen, dass die Funktionswörter in diesen Positionen den Beginn eines neuen Teilsatzes signalisieren und somit die Struktur des Gesamtsatzes durchsichtiger machen.

5.2 Segmentale Information und Wortkategorisierung

Neben kontextuellen *Markern*, die eine syntaktische Umgebung schaffen, die für eine bestimmte Wortkategorie obligatorisch ist, gibt es auch Hinweise im segmentalen Aufbau einer Wortform selbst, die Schlüsse auf seine Wortklassen zulassen. So weisen im Englischen – gerade in kindgerichteter Sprache – Nomen im allgemeinen mehr Silben auf als Verben (Cassidy & Kelly, 1991). Auch über diese Korrelationen scheinen Sprecher des Englischen intuitives Wissen zu haben. Bei der Aufgabe, zu entscheiden ob homophone Nomen und Verben des Englischen häufiger als Nomen oder als Verb verwendet werden, zeigte sich ein deutlicher Einfluss der Silbenzahl des Wortes: bei dreisilbigen Wörtern wurde der Nomengebrauch als frequenter eingeschätzt, bei einsilbigen Wörtern dagegen der Verbgebrauch (Kelly, 1992). In die gleiche Richtung gehen Befunde mit einer anderen Aufgabenstellung von Cassidy und Kelly (1991). Sie gaben ihren Probanden Pseudowörter unterschiedlicher Länge vor. Die Probanden sollten Sätze konstruieren, in denen jeweils eins dieser Pseudowörter verwendet wurde. Die Ergebnisse zeigten wiederum einen Einfluss der Silbenzahl auf den Kontext, der für das Pseudowort konstruiert wurde: je mehr Silben die Pseudowörter aufwiesen, desto geringer war die Wahrscheinlichkeit, dass es als Verb verwendet wurde.

Auch Computersimulationen zeigen die potentielle Nutzbarkeit segmentaler Informationen für die Klassifikation von Verben und Nomen. Durieux und Gillis (2001) testeten, ob die Qualität von Vokalen und Konsonanten sowie die Wortlänge gemessen in der Anzahl der Phoneme potentielle Merkmale sind, die eine Unterscheidung zwischen Verben und Nomen zulassen. Sie untersuchten die Relevanz von Merkmalen die von Kelly (1996) zur Differenzierung von englischen Nomen und Verben

vorgeschlagen wurden, nämlich dass englische Nomen mehr tiefe Vokale und mehr Nasale enthalten als Verben und insgesamt eine höhere Phonemzahl aufweisen. Wie schon bei den Betonungsmerkmalen (vgl. Abschnitt 3.2.1) war das System unter Beachtung dieser Merkmale relativ erfolgreich in der korrekten Klassifikation von Nomen. So wurden anhand des Merkmals Vokalhöhe 91% der Nomen korrekt klassifiziert gegenüber 34% der Verben. Mit dem Merkmal Konsonantenqualität wurden 87% der Nomen korrekt ihrer Wortklasse zugeordnet, aber nur 33% der Verben. Bei Berücksichtigung der Phonemzahl wurden 94% der Nomen und 10% der Verben korrekt klassifiziert. Damit unterschieden sich diese Merkmale nicht wesentlich von der Realibilität von Betonung allein als wortklassendifferenzierendes Merkmal.

In einer letzten Simulation verwendeten die Autoren das – wie sie es nennen – phonologische Rohmaterial des Wortaufbaus, d. h. einfache Phonemketten, die über die einzelnen Segmente hinaus lediglich Information über die Silbengrenzen enthielten. In diesem Fall erhöhte sich vor allen Dingen die Rate der korrekten Wortklassenzuweisungen für die Verben. Sie lag bei 75% während die Nomen mit 79% schlechter als bei den isolierten Klassen von Merkmalen klassifiziert wurden. Mit einem Gesamtanteil von 76% korrekt zugewiesenen Nomen und Verben wurde dieser Input jedoch am erfolgreichsten kategorisiert. Allerdings zeigt sich auch in der Analyse einer weiteren Sprache, dass ein System, das auf gesamten Phonemketten einschließlich Information über Silbengrenzen operiert, erfolgreicher Wortklassenzuordnungen treffen kann, als ein System, das auf klassifizierten Merkmalen basiert. In einer Wiederholung der Simulationen mit einem niederländischen Korpus zeigten sich genau die gleichen Muster: bei der Verwendung von Vokalhöhe, Konsonantenqualität und Segmentzahl wurden mehr Nomen korrekt klassifiziert als Verben. Dieser Unterschied verschwand bei der Verwendung der Phonemketten als Input, wobei der Gesamtanteil der korrekt klassifizierten Nomen und Verben bei 82% war.

Aus diesen Befunden ist zu schließen, dass die einfachen Phonemketten in Verbindung mit den Silbengrenzen Informationen enthalten, die dem System eine Klassifizierung erlauben, die in einer Mehrzahl der Fälle zu einer korrekten Entscheidung führen. Welcher Art diese Informationen sind ist der Arbeit nicht zu entnehmen. Da sich das gleiche Muster in zwei verschiedenen – wenn auch phonologisch relativ ähnlichen - Sprachen findet, ist zu vermuten, dass Korrelationen

zwischen dem segmentalen Aufbau eines Wortes und seiner grammatischen Klasse nicht unbedingt eine ideosynkratische Eigenschaft einzelner Sprachen ist, sondern dass diese in den verschiedensten Sprachen zu finden sein könnten.

MacWhinney, Leinbach, Taraban und McDonald (1989) untersuchten einen anderen Typus von Wortklasseninformation, nämlich die Subkategorisierung in verschiedene Genusklassen innerhalb von Nomen des Deutschen. Die Autoren gehen von Merkmalen aus, die laut Untersuchungen von Köpcke und Zubin (Köpcke, 1982; Köpcke & Zubin, 1984) mit dem Genus eines Nomens korrelieren. Hierbei handelt es sich um eine ganze Reihe verschiedener phonologischer, morphologischer und semantischer Merkmale. In einer ersten Simulation erhielt der Algorithmus als Trainingsset für jedes Nomen die Spezifizierung von insgesamt 18 verschiedenen Merkmalen sowie verschiedenen Kasusformen. Es zeigte sich, dass der Algorithmus anhand dieser Information in der Lage war, bei ca. 60% vorher nicht präsentierter Nomen, den Artikel korrekt vorherzusagen. Interessanterweise berichten die Autoren, von noch höheren Trefferquoten in einer anderen Bedingung, in der als Input lediglich die phonologische Form des Nomens einschließlich fünf verschiedener semantischer Merkmale verwendet wurde. Diese Beobachtung stimmt mit den Befunden von Durieux und Gillis gut überein, die ja ebenfalls fanden, dass die Vorgabe kategorisierter Merkmale dem System weniger nützt als die reine phonologische Form. Demnach ist also allein aus der Analyse der phonologischen Form in Verbindung mit Wissen über semantische Merkmale eine überzufällig häufig korrekte Genuszuweisung möglich. Leider führten die Autoren keine Simulation ohne die Vorgabe semantischer Merkmale durch, deren Ergebnisse für unsere Fragestellung äußerst interessant wären.

5.3 Möglichkeiten und Grenzen der Nutzung segmental-sequentieller Information

Wie vor allen Dingen die Computersimulationen gezeigt haben, enthält menschliche Sprache eine Vielzahl von distributionellen Hinweisen, die ein Verarbeitungs-Algorithmus erfolgreich für die Segmentierung von Äußerungen ausnutzen kann. Insgesamt betrachtet scheint auch das menschliche Sprachverarbeitungssystem in der Lage zu sein, die unterschiedlichsten Hinweise für eine strukturelle Analyse des sprachlichen Inputs auszunutzen. Wie wir gesehen haben, gilt dies für die Ausnutzung sowohl akustischer als auch distributioneller Eigenschaften des sprachlichen Inputs, die für Segmentierungsprozesse auf lexikalischer Ebene, für syntaktische Klassifizierungsprozesse und auch das Erkennen anderer syntaktischer Regularitäten herangezogen werden. Diese Befunde weisen darauf hin, dass das menschliche Verarbeitungssystem offensichtlich in der Lage ist, prinzipiell jegliche Information, die in einer Korrelation mit strukturellen Eigenschaften von Sprache stehen, auszunutzen. Allerdings scheint diese Fähigkeit auch mit einzelsprachlichen Eigenschaften der jeweiligen Muttersprache zu schwanken. Ein Beispiel dafür sind die Befunde bei niederländischen und französischen Sprechern, die im Gegensatz zu finnischen Muttersprachlern keine Evidenz dafür zeigen, dass sie die in ihrer Sprache nicht vorkommende Vokalharmonie ausnutzen können. Wie schon gesagt, lässt sich nicht entscheiden, ob diese Unterschiede zwischen den erwachsenen Sprechern darauf zurückzuführen sind, dass die finnischen Sprecher aufgrund der Tatsache, dass ihre Sprache Vokalharmonie aufweist, gelernt haben, dieses Merkmal zu beachten und es in der Analyse gesprochener Sprache einzusetzen. Die Alternative, um diesen Unterschied zu erklären, bestünde in der Annahme, dass niederländische und französische Sprecher es verlernt haben, diese Merkmale zu beachten, da sie für ihre Sprache nicht relevant sind. Die zweite Alternative ist kompatibel mit den Befunden zur Veränderung in der Lautdiskrimination, wo die Entwicklung von einzelsprachunabhängigen Fähigkeiten zu einer Einschränkung auf die muttersprachlich relevanten Dimensionen zu beobachten ist (vgl. Abschnitt 6.2.1). Entsprechend könnte man vermuten, dass zunächst alle Säuglinge auf akustische Eigenschaften wie Vokalharmonie unabhängig von ihrem sprachlichen

Hintergrund auf gleiche Weise positiv reagieren, und dass in Abhängigkeit von den Merkmalen der zu erwerbenden Sprache, diese Fähigkeit erhalten bleibt oder verloren geht.

Ebenso wie im Bereich der prosodischen Merkmale muss man sich die Frage stellen, ob es im Bereich der nicht-prosodischen distributionellen Information überhaupt Hinweise geben kann, die einem rein datengetriebenen Verarbeitungsmechanismus zugänglich sind bzw. von einem solchen ausgenutzt werden können. Dies lässt sich vielleicht für zwei der angesprochenen Hinweistypen positiv beantworten, nämlich den Harmonieprozessen und der Ausnutzung häufiger Input-Muster.

Durch phonologische Harmonieprozesse werden Sequenzen geschaffen, die in einem akustischen Merkmal übereinstimmen. Möglicherweise führen diese akustischen Übereinstimmungen bzw. Brüche dieser Übereinstimmungen an der Grenze zu einem Segment, das den Harmonieprozessen nicht mehr unterliegt, bereits auf einer grundlegenden Wahrnehmungsebene zur Segmentierung des kontinuierlichen Signals. Dass die Nutzung dieser Informationsquelle zumindest bei Erwachsenen von der sprachlichen Erfahrung beeinflusst wird, wurde bereits angesprochen.

Auch die Feststellung von häufigen Mustern im sprachlichen Signal und die Berechnung von Übergangswahrscheinlichkeiten kann als rein datengetriebener Prozess bezeichnet werden, da lediglich ein Berechnungsmechanismus für die Häufigkeit distributioneller Muster und eine ausreichende Menge von Inputinformationen erforderlich sind. Wie schon erwähnt, ist bislang jedoch lediglich erwiesen, dass das menschliche Verarbeitungssystem im Prinzip über diese Analysemechanismen verfügt. Ob dieses System auch in der Lage ist, distributionelle Information in einem sehr viel variableren Korpus natürlicher Sprache zu extrahieren, in der die Information weniger konzentriert geboten wird, als in der experimentellen Situation, ist bislang eine offene Frage.

Die übrigen diskutierten Hinweise erfordern bereits Wissen über Einheiten auf der Ebene, zu deren Grenzfindung sie herangezogen werden können. Um positionsbezogene allophonische Varianten als Markierung von Wortgrenzen heranziehen zu können, ist Wissen darüber notwendig, dass diese speziell an der Grenze eines Wortes vorkommen. Zu dieser Information kann das System jedoch nur gelangen, wenn es Wortgrenzen bereits erkennen kann. Auch um phonotaktische Restriktionen als

Grenzsignal benutzen zu können, muss das System über Wissen verfügen, welche phonotaktischen Folgen erlaubt sind. Um dieses zu lernen, ist eine vorherige Erkennung der hier relevanten Grenzen notwendig. Für den vermuteten Einsatz lexikalischen Wissens bei der Segmentierung beantwortet sich die Frage von selbst negativ. Dies spricht natürlich nicht dagegen, dass der erwachsene Sprecher diese Hinweise in der Sprachverarbeitung für die effiziente schnelle Strukturierung des Inputs nutzt. Schließlich verfügt der Erwachsene bereits über das sprachliche Wissen, das für die Nutzung dieser Hinweise relevant ist.

Anders sieht es beim Säugling aus. Hier ist zu erwarten, dass die rein datengetriebenen ausnutzbaren Hinweise früher genutzt werden können, als solche, die bereit vorhandenes sprachliches Wissen voraussetzen. Es ist zu erwarten, dass Säuglinge, wie im Bereich der prosodischen Merkmale, die relevanten segmental-sequentiellen Hinweise in ihrem sprachlichen Input sehr früh wahrnehmen können und sie in die strukturelle Analyse ihres sprachlichen Inputs einbeziehen, sobald sie ihnen zur Verfügung stehen. Dieser Annahme wird im folgenden Kapitel weiter nachgegangen.

6 Der segmental-distributionelle Erwerbsstrang

In diesem Kapitel soll dafür argumentiert werden, dass neben prosodischer Information bereits im frühen Spracherwerb zunehmendes Wissen über segmentale und distributionelle Eigenschaften, d.h. Wissen über mögliche Kookurrenzen bzw. die Frequenz von Kookurrenzen auf Phonem- und Silbenebene, später auch lexikalisches Wissen für Strukturierungsprozesse des sprachlichen Inputs herangezogen werden. Dabei steht die Annahme im Vordergrund, dass dieses Wissen sowohl zur Segmentierung des Inputs in linguistisch relevante Einheiten wie Wörter und Phrasen herangezogen werden kann, als auch für eine syntaktische Klassifizierung der so gewonnenen Einheiten. Analog zum Bereich des prosodischen Einstiegs in den Syntaxerwerb sollen hierbei die folgenden Fragen berücksichtigt werden:

- ?? Ab welchem Alter sind Säuglinge sensitiv für die in diesem Zusammenhang relevanten Segmente, d.h. Lautsegmente und Lautsequenzen?
- ?? Inwieweit und in welchem Alter verfügen Säuglingen über Mechanismen der Distributionsanalyse?
- ?? Ab welchem Alter verfügen Säuglinge über sprachspezifisches distributionelles Wissen und beziehen dieses Wissen in die Analyse sprachlicher Reize und den Aufbau sprachlichen Wissens auf lexikalischer und syntaktischer ein?
- ?? Inwieweit kann der segmental-distributionelle Strang Lücken des prosodischen Strangs füllen?

6.1 Voraussetzungen für den Erwerb segmental-distributionellen Wissens

6.1.1 Lautdiskrimination

Die Frage nach der Fähigkeit von Säuglingen, Phoneme zu diskriminieren ist der Ursprung der jüngeren Forschung zum frühen Spracherwerb. In einer bis heute richtungsweisenden Arbeit untersuchten Eimas, Siqueland, Jusczyk und Vigorito (1971) bei ein und vier Monate alten englischen Säuglingen die Fähigkeit, den Stimmhaftigkeitskontrast zwischen den Silben [ba] und [pa] zu diskriminieren. Sie fanden, dass bereits die jüngsten Kinder der Studie diesen Kontrast unterscheiden konnten, und zwar entlang der gleichen Kategoriegrenzen, wie dies bei erwachsenen Sprechern des Englischen der Fall ist. Dieser Befund löste eine ganze Welle weiterer Studien aus, die die Diskriminationsfähigkeit bei Säuglingen für die unterschiedlichsten phonologischen Kontraste belegen. Die Fähigkeit, Kontraste im Artikulationsort von stimmhaften Plosiven wahrzunehmen, ist bereits während der ersten Lebensmonate vorhanden (Eimas, 1974; Jusczyk, 1977; Jusczyk & Thompson, 1978; Moffit, 1971; Morse, 1972) und scheint auch schon zum Repertoire Neugeborener zu gehören (Bertoncini, Bijeljic-Babic, Blumstein & Mehler, 1987). Auch innerhalb anderer Artikulationsartklassen können Säuglinge bereits früh Artikulationsorte differenzieren. Gezeigt wurde dies für verschiedene Liquide (Eimas, 1975), für verschiedene Gleitlaute (Jusczyk, Copan & Thompson, 1978) und verschiedene Nasale (Eimas & Miller, 1980).

Auch können bereits während der ersten Lebensmonate verschiedene Artikulationsarten differenziert werden: so können Verschlusslaute von Gleitlauten unterschieden werden (Eimas & Miller, 1980; Hillenbrand, Minifie & Edwards, 1979; Miller & Eimas, 1983) und orale von nasalen Konsonaten (Eimas & Miller, 1980). Weniger einheitlich sind die Befunde für die Diskrimination verschiedener Frikative, bei denen sich widersprüchliche Ergebnisse während des ersten Lebensjahres gezeigt haben (Eilers, 1977; Eilers, Wilson & Moore, 1977; Levitt, Jusczyk, Murray & Garden, 1988). Ein einheitliches Muster zeigt sich dagegen wieder bei Vokalen. Alle untersuchten vokalischen Kontraste konnten während der ersten Lebensmonate

diskriminiert werden (Trehub, 1973; Swoboda, Morse & Leavitt, 1976; Kuhl, 1983). Die gefundenen Unterscheidungsfähigkeiten sind nicht auf den Silbenanlaut oder den Beginn einer Äußerung beschränkt (Jusczyk, 1977; Jusczyk, Copan & Thompson, 1978; Jusczyk & Thompson, 1978).

Offensichtlich bilden auditive Erfahrungen mit den akustischen Mustern, die diesen Kontrasten entsprechen, keine Voraussetzung für die Fähigkeit, sie zu unterscheiden. Dafür spricht zum einen die Beobachtung, dass bereits Neugeborene bestimmte Kontraste differenzieren können (Bertoncini et al., 1988), zum anderen aber auch der vielfach replizierte Befund, dass Säuglinge auch Kontraste diskriminieren können, die in der Umgebungssprache nicht vorkommen bzw. phonologisch nicht relevant sind. So können amerikanische Säuglinge die negativen und positiven Stimmeinsatzzeiten des Thai differenzieren, obwohl das Englische keine Konsonanten mit negativen Stimmeinsatzzeiten aufweist (Aslin, Pisoni, Hennessy & Perey, 1981). Auch innerhalb der Artikulationsort-Kontraste zeigen sich analoge Ergebnisse. Sechs Monate alte amerikanische Säuglinge sind in der Lage einen dental/retroflex Kontrast aus dem Hindi und einen velar/uvular Kontrast aus dem Nthalkapmx zu unterscheiden, die im Englischen nicht vorkommen (Werker & Tees, 1984). Ein- bis viermonatige kanadische Säuglinge mit Englisch als Zielsprache können nasale von oralen Vokalen unterscheiden (Trehub, 1976). Viermonatige englische Kinder sind für den deutschen Kontrast zwischen vorderen gerundeten und hinteren gerundeten Vokalen ([u] vs. [y]) sensitiv (Polka & Werker, 1994). Ein bis vier Monate alte Kikuyu-Kinder können den Stimmhaftigkeitskontrast zwischen den Verschlusslauten [b] und [p] diskriminieren, obwohl er in ihrer Sprache phonologisch nicht relevant ist (Streeter, 1976). Vier bis sechs Monate alte spanische Säuglinge differenzieren die Stimmhaftigkeitskategorien des Englischen, obwohl im Englischen die Kategoriegrenze zwischen stimmhaften und stimmlosen labialen Verschlusslauten einen anderen *Voice-Onset-Time*-Wert hat als im Spanischen (Lasky, Syrdal-Lasky & Klein, 1975).

Zusammengefasst deuten die Befunde daraufhin, dass Säuglinge in den ersten Lebensmonaten bereits über weitreichende Fähigkeiten verfügen, die unterschiedlichsten Lautkontraste zu differenzieren. Sprachliche Erfahrung scheint hierbei keine Rolle zu spielen. Diese Befunde stützen die Annahme, dass Säuglinge mit der Fähigkeit geboren werden, alle potentiell in irgendeiner Sprache der Welt

auftretenden lautlichen Kontraste wahrzunehmen. Da sich Sprachen in ihren phonologischen Systemen deutlich unterscheiden, ist diese Offenheit für die unterschiedlichsten akustischen Kategorien eine entscheidende Voraussetzung für einen schnellen und effizienten Erwerb des lautlichen Inventars und des lautlichen Systems der Zielsprache.

6.1.2 Bildung von Repräsentationen

Eine Voraussetzung für den Erwerb phonotaktischer Regularitäten ist, dass relevante Einheiten aus dem Sprachstrom extrahiert werden können und zumindest eine kurzfristige Repräsentation dieser Sequenzen verfügbar ist, auf Grundlage derer Berechnungen über ihre Bestandteile und deren Kombination möglich sind. Für den Bereich phonotaktischer Regularitäten ist zunächst anzunehmen, dass in diesem Zusammenhang Silben eine bedeutsame Rolle spielen. Es ist also zu fragen, ob und wann während des frühen Spracherwerbs eine Sensitivität für Silben und eine Fähigkeit, Lautsequenzen zumindest über einen kurzen Zeitraum mental zu repräsentieren, vorhanden sind.

Die im vorangegangenen Abschnitt dargestellten Befunde, nach denen Säuglinge in der Lage sind, phonetische Kontraste wahrzunehmen, die phonologisch Einzelphonemen zugeordnet werden, lassen sich nicht unbedingt dahingehend interpretieren, dass Säuglinge einzelne Laute bereits als aus dem Kontext herauslösbare Segmente identifizieren, da in allen Untersuchungen Silben und - zumindest im Falle der Konsonanten – nicht einzelne Laute als Stimuli präsentiert wurden. Somit könnten die Befunde auch als Evidenz für eine vorhandene Diskriminationsfähigkeit für gesamte Silben interpretiert werden. Es spricht einiges dafür, dass die Segmentierung von Sprache in Silben eine natürlichere Fähigkeit darstellt, als die Segmentierung in einzelne Laute. Hinweise darauf ergeben sich aus den bereits in Abschnitt 4.1.2.1 dargestellten Befunden, nach denen 4 Tage alte Säuglinge sprachliche Reize mit unterschiedlicher Silbenzahl diskriminieren aber nicht sprachliche Reize, die sich nur in der Anzahl der Lautsegmente unterscheiden (Bijeljic-Babic et al., 1993).

Untersuchungen mit älteren Kinder zeigen, dass diese noch bis ins Schulalter größere Probleme bei Aufgaben haben, die eine Segmentierung sprachlicher Reize in

einzelne Laute erfordern, als bei Aufgaben, die eine Silbensegmentation erfordern (z.B. Liberman, Shankweiler, Fischer & Carter, 1974; Treiman & Baron, 1981; Treiman & Zukowski, 1991). Diese Diskrepanz zwischen Lautsegmenten und Silben zeigt sich sogar noch bei erwachsenen Analphabeten (Morais, Bertelson, Cary & Alegria 1986; Morais, Cary, Alegria, Bertelson, 1979) und Schreibern nicht-orthographischer Schriftsysteme (Read, Zhang, Nie & Ding, 1986), so dass die Vermutung nahe liegt, dass die Silbe eine natürlichere, der Wahrnehmung besser zugängliche strukturelle Einheit darstellt als das Lautsegment, dessen Segmentierung möglicherweise eine rein metasprachliche Fähigkeit darstellt, die eng mit dem Erwerb eines orthographischen Schriftsystems zusammenhängt (vgl. Höhle & Weissenborn, 2000).

Zwar wurde die oben erwähnt angeborene Sensitivität gegenüber der Silbenzahl von sprachlichen Stimuli in einer weiteren Studie repliziert (Bertoncini et al., 1995), bislang ist jedoch nichts darüber bekannt, an welchem akustischen Korrelat sie festzumachen ist. Es ist möglich, dass die rhythmische Sensitivität von Säuglingen zu den beobachteten Reaktionen auf die Veränderung der Silbenzahl führt, da eine Veränderung der Silbenzahl einer Sequenz auch zur Veränderung des rhythmischen Musters führt, was bei der Veränderung der Menge der Lautsegmente nicht der Fall sein muss. Es könnte auch sein, dass die Menge der vokalischen Anteile im Signal für die Diskrimination verschiedener Silbenzahlen relevant ist.

Eine zumindest kurzfristige Speicherung von Lautsequenzen und die Fähigkeit, diese auch unter veränderten Inputbedingungen wiederzuerkennen, belegen die Untersuchungsergebnisse von Jusczyk und Aslin (1995). In dieser Studie wurden siebeneinhalb Monate alten Kindern zunächst isolierte einsilbige Nomen mehrfach präsentiert. Anschließend hörten die Kinder kurze Textpassagen, in denen diese Nomen wieder enthalten waren oder nicht. Es zeigte sich eine Präferenz für die Textpassagen, die die zuvor gehörten Nomen enthielten. Dieses Ergebnis spricht dafür, dass die Kinder eine zumindest kurzfristige Repräsentation der Lautsequenz der Nomen aufbauten und die in den Texten auftauchenden weiteren Token der Nomen auf diese Repräsentation abbilden konnten, obwohl diese zunächst aus dem Kontext, in dem sie auftraten herausgelöst werden mussten. Wie ein weiteres Experiment dieser Untersuchungsreihe zeigte, lässt sich dieser Wiedererkennungseffekt auch bei umgekehrter Darbietung der Stimuli erzielen, d. h. wenn den Kindern zunächst die Textpassagen präsentiert werden,

reagieren sie bei einer anschließenden isolierten Darbietung der Nomen stärker auf diejenigen, die zuvor in den Textpassagen vorkamen. Dies setzt voraus, dass die Kinder bereits in der Lage sind, die entsprechenden Lautsequenzen aus den kontinuierlichen Textpassagen zu extrahieren. Auf Grundlage welcher Segmentierungsmechanismen dies passiert, ist bislang unklar. Es könnte sich um eine silbische Repräsentation handeln, wobei die mehrfache Wiederholung der Nomen in den Texten ihre Repräsentation als eigenständige Lautsequenz wahrscheinlich unterstützt.

Neben der Fähigkeit, die relevanten Segmente im sprachlichen Input zu erkennen, ist eine weitere Voraussetzung für das Lernen distributioneller Eigenschaften der Sprache, dass eine Repräsentation dieser Segmente gebildet wird und weitere, im Input auftretende Realisierungen dieser Segmente dieser mentalen Repräsentation zugeordnet werden können.

Einige Studien gingen der Frage nach, über welchen Zeitraum mentale Repräsentationen sprachlicher Reize verfügbar bleiben. Übereinstimmend zeigen diese Untersuchungen, dass sich auch dann noch Wiedererkennungseffekte für Wörter nachweisen lassen, wenn die Familiarisierung und die Testphase in einem Experiment nicht direkt nacheinander durchgeführt werden, sondern mit zeitlichen Abständen von bis zu zwei Wochen (Jusczyk & Hohne, 1997; Houston & Jusczyk, 2001). Dies zeigt, dass die nach nur wenigen Präsentationen der Wörter aufgebaute Repräsentation durchaus schon über einen längeren Zeitraum verfügbar ist, also eine Langzeitrepräsentation gebildet worden ist.

Eine wesentliche Voraussetzung dafür, dass weitere Realisierungen im Input diesen mentalen Repräsentationen zugeordnet werden können, ist die Fähigkeit von bestimmten Merkmalen der phonetischen Realisierung zu abstrahieren. Auch dies scheint schon in sehr frühem Alter möglich zu sein: so zeigt sich, dass Säuglinge bereits in ihren ersten Lebensmonaten in Lautdiskriminationsaufgaben über verschiedene inter- und intraindividuelle Veränderungen der Signale bei verschiedenen Sprechern, Stimmlagen und Sprechgeschwindigkeiten abstrahieren können ((Jusczyk, Pisoni & Mullennix, 1992; Kuhl, 1979, 1983; Miller & Eimas, 1983). Zudem zeigte sich, dass Säuglinge im Alter von 7 Monaten familiarisierte Wörter auch dann wiedererkennen, wenn die isolierten Wörter und die Textpassagen nicht vom gleichen Sprecher gesprochen wurden (Houston & Jusczyk, 2000, 2001). Diese Befunde zeigen

deutlich, dass die gebildeten Repräsentationen nicht alle akustischen Eigenschaften des Signals enthalten, bzw. ein Signal auch dann einer mentalen Repräsentation zugeordnet wird, wenn sich nicht alle akustischen Eigenschaften des Signals in dieser Repräsentation wiederfinden. Segmental scheinen die Repräsentationen jedoch schon relativ gut spezifiziert zu sein, denn es finden sich keine Wiedererkennungseffekte, wenn sich die familiarisierten Stimuli von den Teststimuli in nur einem phonetischen Merkmal unterscheiden (z. B. *cup* – *cut*; Jusczyk & Aslin, 1995).

Eine eigene Untersuchung zur Wiedererkennung von Funktionswörtern (Höhle & Weissenborn, in press) liefert zusätzliche Evidenz, dass Säuglinge im Alter von 7 Monaten bereits Repräsentationen von Lautsequenzen aufbauen und unterschiedliche Token dieser Sequenzen auch unter deutlich veränderten akustischen Bedingungen ein und derselben Repräsentation zuordnen können. Die Vorgehensweise in diesem Experiment war vollkommen analog zu der zuvor dargestellten Untersuchung von Jusczyk und Aslin, der einzige wesentliche Unterschied bestand darin, dass statt einsilbiger Nomen einsilbige Elemente der geschlossenen Klasse verwendet wurde. Diese wurden in den konstruierten Textpassagen in unbetonten Positionen verwendet, so dass eine größere Varianz in den akustischen Merkmalen der einzelnen für die Familiarisierung eingesprochenen Token und den innerhalb der Textpassage auftretenden Token zu erwarten ist.

6.1.2.1 Experiment 5: Wiedererkennung unbetonter Elemente der geschlossenen Klasse im Kontext

Probanden

Analog zu der Untersuchung von Jusczyk und Aslin (1995) wurden zwei Gruppen von Kindern im Alter zwischen 6 und 9 Monaten in die Untersuchung einbezogen. Die ältere Gruppe bestand aus 28 termgeborenen Kindern, davon 15 Mädchen und 13 Jungen. Es nahmen nur Kinder teil, die in einem einsprachigen Elternhaus aufwuchsen und zu denen ausschließlich Deutsch gesprochen wurde. Das mittlere Alter in dieser Gruppe betrug 8 Monate und einen Tag und reichte von 7 Monaten und 14 Tagen bis zu 8 Monaten und 30 Tagen. Um auf die Zahl von 28 verwertbaren Datensätzen zu kommen, wurden 11 weitere Kinder getestet, deren Ergebnisse aufgrund von

Experimentalabbrüchen oder zu kurzen Orientierungszeiten nicht in die Datenanalyse eingingen.

Auch die jüngere Altersgruppe umfasste 28 Kinder, von denen 13 Mädchen und 15 Jungen waren. Es wurden wiederum nur Kinder mit einsprachigem Elternhaus in die Studie einbezogen. Die Kinder hatten ein durchschnittliches Alter von 6 Monaten und 12 Tagen bei einer Spanne von 6 Monaten und 3 Tagen bis zu 6 Monaten und 26 Tagen. Drei zusätzliche Kinder wurden getestet, mit denen das Experiment nicht beendet werden konnte oder die zu kurze durchschnittliche Orientierungszeiten von unter 3 s aufwiesen und daher nicht in die Datenanalyse eingingen.

Untersuchungsmaterial

Als Wortmaterial dienten vier Elemente der geschlossenen Klasse: der definite Artikel bzw. die Konjunktion *das*, das Pronomen *sein* sowie die beiden Präpositionen *bis* und *von*. Bis auf den Faktor der Betonung sollte eine möglichst hohe Vergleichbarkeit der Studie zur Untersuchung von Jusczyk und Aslin erzielt werden, daher wurden ausschließlich Wörter benutzt, die aus einer geschlossenen Silbe bestanden. Die Auswahl wurde zudem dadurch eingeschränkt, dass die vier Wörter sowohl im Anlaut als auch im Vokal keine Übereinstimmungen aufweisen sollten. Für jedes Wort wurde eine Textpassage konstruiert, in der das kritische Wort in jedem Satz genau einmal vorkam. Dabei wurde darauf geachtet, dass keins der Testwörter auch in einer anderen Passage des Sets vorkam, d.h. beispielsweise wurde die Artikelform *das* in keiner der Textpassagen für die übrigen drei Wörter verwendet. Die einzelnen Textpassagen bestanden jeweils aus sechs syntaktisch einfachen Sätzen. Die Position des kritischen Wortes wurde zwischen den Sätzen variiert.

Das Untersuchungsmaterial wurde von einer Muttersprachlerin des Deutschen gesprochen. Sie erhielt die Anweisung, die Texte und Wörter mit variierender Stimmlage, so als ob sie zu einem kleinen Kind sprechen würde, zu lesen. Die einzelnen Wörter wurden von ihr in ca. 30 verschiedenen Varianten mit der expliziten Anweisung, diese prosodisch etwas unterschiedlich zu realisieren, gesprochen. Das Material wurde digitalisiert und als Datei abgespeichert.

Für die Familiarisierung wurde für jedes Wort eine akustische Datei hergestellt, die verschiedene Varianten des kritischen Wortes enthielt, wobei jeweils eine Pause von

600 ms zwischen die einzelnen Wiederholungen des Wortes eingefügt wurde. Diese Dateien hatten eine Länge von ca. 30 Sekunden und enthielten 28 Realisierungen des Zielwortes. Auch für jede Textpassage wurde eine eigene Datei angelegt. Die Texte hatten jeweils eine Dauer von 18 bis 20 Sekunden.

Um sicherzustellen, dass es tatsächlich deutliche akustische Unterschiede zwischen den isoliert eingesprochenen Wörtern und denen in den Textpassagen gab, wurde eine Analyse der Wortlänge und der Grundfrequenz für alle Token der Wörter durchgeführt. Die isoliert eingesprochenen Token wiesen eine durchschnittliche Länge von 494 ms auf, während die kontextuell eingebetteten Token durchschnittlich 208 ms lang waren. Diese Werte belegen die Annahme deutlicher akustischer Unterschiede zwischen den isolierten und den kontextuell eingebetteten Token, und reflektieren die Tatsache, dass die kritischen Wörter in den Textpassagen in unbetonten Positionen auftreten.

Untersuchungsdurchführung

Das Experiment wurde anhand des Headturn Preference Paradigmas mit Familiarisierungsphase durchgeführt. Während der Familiarisierung wurden jedem Kind zwei der vier Wörter des gesamten Sets präsentiert, und zwar so lange bis das Kind jedes der beiden Wörter für mindestens 30 Sekunden gehört hatte. Die Präsentation war abhängig vom Verhalten des Kindes: nur so lange das Kind den Kopf in Richtung des Lautsprechers gedreht hielt, wurde der sprachliche Reiz präsentiert. Drehte das Kind den Kopf länger als 2 Sekunden weg, wurde der Durchgang gestoppt und zum nächsten übergegangen. Aufgrund dieses Vorgehens konnte das Kind die notwendige Familiarisierungszeit für die beiden Wörter entweder innerhalb eines einzigen Familiarisierungsdurchgangs pro Wort erreichen oder in mehreren Durchgängen aufsummieren. Zwischen den Kindern wurde variiert welche zwei der vier Wörter sie während der Familiarisierung hörten.

In der Testphase des Experiments wurden allen Kindern alle vier Textpassagen für alle vier kritischen Wörter präsentiert. Durch die Variation der Familiarisierungssitems bei Konstanthaltung des Materials in der Testphase, wird sichergestellt, dass gefundene Unterschiede zwischen den Texten mit familiarisierten Wörtern und Texten ohne familiarisierte Wörter nicht auf generellen Hörpräferenzen für einen Teil der Texte beruhen. Die Testphase bestand aus vier Präsentationsblöcken, die

jeweils alle vier Textpassagen enthielten. Zwischen den Blöcken war die Reihenfolge der Texte verschieden.

Ergebnisse und Diskussion der Ergebnisse

Die 8 Monate alten Kinder zeigten einen signifikanten Familiarisierungseffekt: mit durchschnittlich 8,34 s orientierten sie sich länger zur Schallquelle, wenn ein Text präsentiert wurde, der ein familiarisiertes Wort enthielt, gegenüber 7,40 s bei Texten, die keins der familiarisierten Wörter enthielten. Nach dem t-Test für abhängige Stichproben war dieser Unterschied signifikant ($t(df = 27) = 2,77$; $p = 0,011$). Einundzwanzig der 28 untersuchten Kinder zeigten diese Präferenz. Bei den Kindern der jüngeren Altersgruppe von 6 Monaten zeigte sich dagegen kein statistisch signifikanter Unterschied in den Orientierungszeiten: hier fanden sich 8,76 s für die Texte mit familiarisierten Wörtern gegenüber 8,62 s für die Texte ohne familiarisierte Wörter ($t(df = 27) = 0,42$; $p = 0,67$). Lediglich 16 der 28 Kinder zeigten längere Orientierungszeiten für die Texte mit familiarisierten Wörtern.

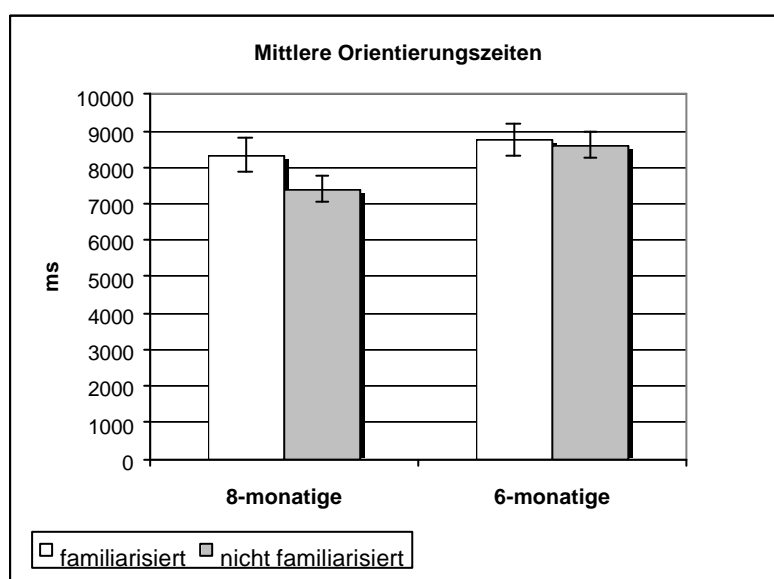


Abbildung 5: Mittlere Orientierungszeiten für Texte mit familiarisierten Wörtern und Texte ohne familiarisierte Wörter

Unsere Ergebnisse zeigen ein zeitlich vollkommen analoges Muster zu den Befunden von Jusczyk und Aslin (1995) für Nomen: ab einem Alter von ca. siebeneinhalb

Monaten zeigen die Kinder längere Orientierungszeiten, wenn ihnen Textpassagen präsentiert werden, die die zuvor gehörten Wörter enthalten gegenüber Textpassagen, die diese Wörter nicht enthalten. 6 Monate alte Säuglinge zeigen diesen Effekt noch nicht. Diese Daten liefern zusätzliche Evidenz, dass Kinder diesen Alters bereits Token verschiedener akustischer Qualitäten auf ein und dieselbe mentale Repräsentation abbilden können. Außerdem – und dies ist ein wichtiger Befund im Rahmen unserer Hypothese, dass die Information von Funktionswörtern bereits früh im Spracherwerb für die Sprachverarbeitung und den Aufbau sprachspezifischen Wissens eingesetzt wird – zeigen die Befunde, dass der Faktor der Unbetontheit keine gravierenden negativen Konsequenzen für die Wahrnehmungs- und Verarbeitungsleistungen bei jüngeren Kindern zu haben scheint. Vergleicht man die hier präsentierten Daten zur Detektion unbetonter Elemente mit denen von Jusczyk und Mitarbeitern zur Detektion von Nomen zeigen sich weitgehend parallele Befunde. Dies spricht gegen die Annahme, dass die relativ späte Produktion von Funktionswörtern mit ihrer geringen phonologischen Salienz im sprachlichen Input der Kinder zu tun haben könnte (z.B. Bates & Goodman, 1999; Gleitman & Wanner, 1982).

Mit den nachgewiesenen Mitteln zur segmentalen Analyse sowie zum Aufbau und zur Speicherung von Repräsentation von Lautsequenzen verfügen die Kinder also schon frühzeitig über eine der Voraussetzungen, die notwendig sind, um die Auftretenshäufigkeit von Lautsegmenten sowie von Silben und deren Kombinationen in ihrem sprachlichen Input zu berechnen. Dass Kindern diese Möglichkeiten zur Distributionsanalyse schon während des ersten Lebensjahres zur Verfügung stehen und welche weiteren Schritte in Richtung des Erwerbs von muttersprachlichen Charakteristika ihnen diese Analysemechanismen eröffnen, soll im folgenden gezeigt werden.

6.1.3 Analysemechanismen für distributionelle Information

Eine der ersten Arbeiten, die sich mit der Frage nach der Sensitivität von Säuglingen gegenüber distributionellen Eigenschaften ihres sprachlichen Inputs beschäftigt, ist die Studie von Goodsitt, Morgan und Kuhl (1993). Sie fanden, dass sieben bis acht Monate alte Säuglinge zwei Silben in einem veränderten Kontext dann besser wiedererkennen,

wenn ihnen diese Silben zuvor mehrfach adjazent in gleicher Reihenfolge präsentiert worden waren, als wenn ihnen die Silben zwar adjazent aber in variierender Reihenfolge geboten worden waren.

In die gleiche Richtung gehen auch Befunde von Morgan und Saffran (1995). Sie fanden bei 9 Monate alten Säuglinge, längere Latenzen bei der Reaktion auf ein Störgeräusch, wenn dieses zwischen zwei Silben auftrat, die vorher mehrfach paarweise in konstanter Reihenfolge präsentiert worden waren, als wenn dieses zwischen zwei Silben auftrat, die vorher in variabler Abfolge dargeboten worden waren.

Eindrucksvolle Belege für die Effizienz der Analysemechanismen, die Säuglingen zur Verfügung stehen, stammen aus einer Arbeit von Saffran, Aslin und Newport (1996, s. a. Saffran, 2001). Für ihre Untersuchung verwendeten sie vier dreisilbige Pseudowörter (z.B. *daropi*, *pakibu*), die 8 Monate alten Säuglingen zunächst als kontinuierlicher Sprachstrom in verschiedenen Reihenfolgen für 2 Minuten präsentiert wurden. Die Reihenfolge der Pseudowörter war so festgelegt, dass jede Folge zweier Pseudowörter in gleicher Häufigkeit auftrat. Auf diese Weise wurde erreicht, dass die Übergangswahrscheinlichkeit der Silben innerhalb eines Pseudowortes 100% betrug und die Übergangswahrscheinlichkeiten zwischen den jeweils letzten und den ersten Silben der verschiedenen Pseudowörter 30% betrug. In der Testphase des Experiments wurden die Präferenzen der Kinder für ein gesamtes Pseudowort gegenüber einem „neuen“ Pseudowort, das aus der letzten Silbe eines Pseudowortes und den beiden ersten Silben eines anderen Pseudowortes der Familiarisierungsphase gebildet worden war (z. B. für obige Beispiele: *pipaki*), überprüft. Es fand sich bei den Kindern eine Präferenz für die „neuen“ Pseudowörter. Aus diesem Neuigkeitseffekt lässt sich ableiten, dass Kinder diesen Alters hochsensitiv für die Übergangswahrscheinlichkeiten zwischen den Silben sind. Sie sind nicht nur in der Lage zu erkennen, ob zwei Silben miteinander vorgekommen sind oder nicht, sie können darüber hinaus die Häufigkeit dieses gemeinsamen Auftretens berechnen (s.a. Johnson & Jusczyk, 2001).

Befunde von Marcus, Vijayan, Rao und Vishton (1999) weisen darauf hin, dass schon sieben Monate alte Säuglinge die Fähigkeit haben, aus der Abfolge bestimmter Elemente im sprachlichen Input Regularitäten abzuleiten und auf neue Elemente zu übertragen. Sie boten ihren Untersuchungsteilnehmern zunächst verschiedene

dreisilbige Sequenzen dar, denen gemeinsam war, dass an den gleichen Positionen jeweils eine Silbe mehrfach vorkam (z.B. *ga-ti-ga*, *li-na-li*). 16 solcher Sequenzen wurden den Kindern je dreimal präsentiert. Anschließend wurden in der eigentlichen Testphase neue Sequenzen geboten, deren Aufbau entweder der familiarisierten Struktur (z.B. *ki-na-ki*) entsprach, oder deren Aufbau eine andere Struktur aufwies (z.B. *na-ki-ki*). Sie fanden, dass die Säuglinge jeweils die Sequenzen präferierten, die die nicht familiarisierte Struktur aufwiesen. Diese Ergebnisse deuten darauf hin, dass die Kinder über Analyse- und Abstraktionsmechanismen verfügen, die eine Grundlage zum Aufbau syntaktischer Regeln darstellen.

Bei etwas älteren Kindern zeigt sich die Effizienz dieser Mechanismen auch unter variableren und damit komplexeren Lern- und Testbedingungen. Gomez und Gerken (1999) präsentierten 12 Monate alten Kindern zunächst 10 Silbensequenzen mit einer Länge zwischen 3 und 6 Silben. Diese Sequenzen waren aus insgesamt 5 verschiedenen Silben konstruiert worden. Die Silbensequenzen waren entsprechend der Regeln einer künstlichen Grammatik erzeugt, so dass die Abfolge der Silben in den Sequenzen zwar regelhaft war, aber nicht in allen Sequenzen identisch. In der eigentlichen Testphase des Experiments wurden den Kindern Silbenfolgen präsentiert, die in der Lernphase noch nicht vorgekommen waren, die jedoch entweder entsprechend des zugrundeliegenden Regelsystems konstruiert worden waren (grammatische Sequenzen) oder nicht (ungrammatische Sequenzen). Die wesentlichen Unterschiede zur Studie von Saffran und Mitarbeitern bestehen also darin, dass die Silbensequenzen der Lernphase weitaus variabler sind und dass die Testsequenzen neue Kombinationen beinhalten, die die Kinder während der Lernphase noch nicht gehört hatten. In verschiedenen Experimenten mit dem *Headturn Preference* Paradigma konnten die Autorinnen zeigen, dass die Kinder nach dreimaliger Präsentation der Lernsequenzen Wissen darüber erworben haben, welche Silben am Anfang und am Ende einer Sequenz vorkommen dürfen und welche paarweisen Kombinationen von Einzelsilben innerhalb der Sequenzen möglich sind. Dass die Fähigkeiten der Kinder über das Lernen von Paarassoziationen hinausgeht, zeigte sich daran, dass die Kinder auch bei Veränderungen des Vokabulars zwischen Lern- und Testsequenzen die grammatischen von den ungrammatischen Sequenzen unterscheiden konnten. Damit scheinen die Kinder über Möglichkeiten zu verfügen, aus der Reihenfolge von Silben auf

zugrundeliegende abstraktere Regularitäten zu schließen und diese Regularitäten auch auf verändertes Vokabular übertragen zu können. Hiermit ist eine wesentliche Voraussetzung zur Erfassung syntaktischer Regularitäten und deren Abstrahierung von einer gegebenen sprachlichen Äußerung gegeben.

6.2 Der Aufbau sprachspezifischen distributionellen Wissens

Im vorherigen Abschnitt wurde gezeigt, dass Säuglinge während ihres ersten Lebensjahres über vielfältige Voraussetzungen zum Erwerb von segmental-distributionellen Merkmale aus ihrer Umgebungssprache verfügen: sie sind in der Lage, einzelne Laute und/oder Silben voneinander zu unterscheiden, sie können Repräsentationen von Lautsequenzen bilden und verschiedene Realisierungen dieser Sequenzen einer Repräsentation zuordnen. Außerdem zeigen sie beachtliche Fähigkeiten, distributionelle Merkmale ihres Inputs zu erkennen. Im folgenden Abschnitt soll der Frage nachgegangen werden, inwieweit sich in diesen Fähigkeiten Effekte der Umgebungssprache zeigen bzw. ab welchem Alter Säuglinge spezifische Eigenschaften ihrer Umgebungssprache erkannt haben.

6.2.1 Der Aufbau des Phoneminventars

Der Aufbau segmental-distributionellen Wissens, d. h. Wissen über das spezifische Inventar von Elementen, deren Auftretenshäufigkeit sowie deren Kombinationsmöglichkeiten in der Zielsprache, lässt sich zunächst in der Domäne der Phonologie beobachten. Hierzu gehören sowohl der Aufbau des segmentalen Inventars der Zielsprache, der eine wesentliche Voraussetzung für die Perzeption und Produktion muttersprachlicher Äußerungen darstellt, als auch der Aufbau von Kenntnissen über phonotaktische Regularitäten.

Wie oben dargelegt, scheint die Fähigkeit, ein universelles Repertoire an phonetischen Kontrasten wahrnehmen und diskriminieren zu können, zu den dem Menschen angeborenen Verarbeitungsmechanismen zu gehören. Wie erwähnt, sind

diese Mechanismen eine Voraussetzung dafür, die Phonemsysteme der verschiedenen Sprachen der Welt erwerben zu können, die sich ja gerade darin unterscheiden, dass unterschiedliche phonetische Kontraste phonologisch genutzt werden. Zahlreiche Untersuchungen haben gezeigt, dass Erwachsene Säuglingen in ihrer Fähigkeit, lautliche Kontraste zu unterscheiden, deutlich unterlegen sind (z.B. Werker, Gilbert, Humphrey & Tees, 1981; Werker & Tees, 1984). Velmehr zeigt sich bei Erwachsenen ein Einfluss der Muttersprache: ihre Diskriminationsleistungen für fremdsprachliche Kontraste sind im allgemeinen schlechter als für muttersprachliche Kontraste (z.B. Flege, 1989; Logan, Lively & Pisoni, 1989; Miyawaki, Strange, Verbrugge, Liberman, Jenkins & Fujimura, 1975; Strange & Jenkins, 1978), wobei allerdings die phonetische Nähe des fremdsprachlichen Kontrasts zu einem in der eigenen Sprache genutzten eine Rolle spielt (Best, 1990; Best, McRoberts & Sithole, 1988). Darüber hinaus zeigt sich bei Erwachsenen innerhalb kontinuierlicher phonetischer Dimensionen eine besonders hohe Sensitivität für phonetische Veränderungen in den Bereichen des Kontinuums, in denen die Grenzen zwischen den Phonemkategorien in der Muttersprache liegen, wobei hier ein Einfluss der muttersprachlichen Kategoriegrenzen festzustellen ist (z.B. Liberman, Harris, Kinney & Lane, 1961; Liberman, Cooper, Shankweiler & Studdert-Kennedy, 1967). Beide Phänomene können als Adaptation des Verarbeitungssystems an die spezifischen Gegebenheiten des muttersprachlichen Systems interpretiert werden, die die Sensitivität für phonologisch relevante Eigenschaften gegenüber weniger relevanten phonetischen Eigenschaften schärft. Damit verfügt das Verarbeitungssystem über ein Instrument, um im Signal phonologisch relevante Merkmale vom durch die akustische Varianz erzeugten Rauschen zu trennen.

Die Entwicklung von der universellen zur sprachlich beeinflussten Wahrnehmung vollzieht sich bereits während des ersten Lebensjahres. So zeigt sich in Befunden von Werker und Tees (1984) ein deutliches Absinken der Fähigkeit zur Diskrimination nicht muttersprachlicher Kontraste in der zweiten Hälfte des ersten Lebensjahres: während 6 bis 8 Monate alte englische Säuglinge noch die volle Diskriminationsfähigkeit für zwei Hindi-Kontraste und zwei Nthlakapmx-Kontraste aufweisen, die im phonologischen System des Englischen nicht vorkommen, zeigen 8 bis 10 Monate alte Säuglinge eine Abschwächung dieser Fähigkeit. Im Alter von 10 bis 12 Monaten zeigten die Kinder wie Erwachsene keine Evidenz mehr für die

Diskrimination der nicht-muttersprachlichen Kontraste. Im gleichen Altersbereich fanden Werker und Lalonde (1988) bei englischen Säuglingen Evidenz für schwächer werdende Diskriminationsleistungen mit einem weiteren Hindi-Kontrast.

Wie Best und Mitarbeiter zeigen konnten (Best et al., 1988; s. a. Best, 1993), finden sich aber nicht für alle nicht-nativen Kontraste Veränderungen der Diskriminationsleistungen während des ersten Lebensjahres, so blieben sie bei englischen Säuglingen beispielsweise für einen Kontrast bei Klicklauten des Zulu erhalten. Best (1993) vermutet, dass die Diskriminationsfähigkeit nur für solche nicht-nativen Kontraste abnimmt, die dem muttersprachlichen phonologischen System ähnlich genug sind, um in diesem auf eine einzige Phonemkategorie abgebildet zu werden. Diese Abbildung auf eine einzige Phonemkategorie führt dann wiederum zu schlechten Diskriminationsleistungen. Keine Veränderungen in der Diskriminationsfähigkeit sollten sich dagegen für solche nicht-nativen Kontraste zeigen, die auf verschiedene Phonemkategorien der Muttersprache abgebildet werden und für solche Kontraste, die weit außerhalb des Phonemsystems der Muttersprache liegen und daher gar nicht auf dieses abgebildet werden.

Auch im Bereich der Vokalwahrnehmung finden sich Einflüsse des muttersprachlichen Inputs. Im Gegensatz zur Diskriminationsfähigkeit bei Konsonanten scheint diese bei Vokalen kontinuierlicher zu sein, so dass phonetisch verschiedene Exemplare derselben Vokalkategorie von Erwachsenen besser diskriminiert werden können als phonetisch verschiedene Realisierungen derselben Konsonatenkategorie (z.B. Pisoni, 1973). Ein Einfluss der muttersprachlichen Organisation der Vokalkategorien zeigt sich allerdings im sogenannten „perzeptuellen Magneteffekt“ (s. z.B. Kuhl, 1993). Dieser Effekt besagt, dass verschiedene Realisierungen eines Vokals schlechter von einer in ihren phonetischen Eigenschaften als sehr prototypisch eingestuften Realisierung diskriminiert werden können, als von einer weniger prototypischen Realisierung. Kuhl (1993) fand, dass bereits Säuglinge im Alter von 6 Monaten diesen Magneteffekt für ihre Muttersprache zeigen. Dass der perzeptuelle Magneteffekt durch die Verhältnisse der Umgebungssprache geprägt wird, zeigt sich daran, dass sowohl Erwachsene als auch 6 Monate alte Säuglinge ihn nicht für nicht-native Vokalkategorien zeigen, d. h. innerhalb nicht-nativer Vokalkategorien sind die Diskriminationsleistungen besser als innerhalb nativer (Kuhl, Williams, Lacerda,

Stevens & Lindblom, 1992; s. a. Kuhl, 1993). Nach diesen Befunden ist davon auszugehen, dass der Magneteffekt nicht auf rein phonetischen Faktoren beruht, sondern die Erfahrung mit der Wahrnehmung der spezifischen phonetischen Eigenschaften der Vokalrealisierungen in der sprachlichen Umgebung voraussetzt.

Genau wie bei den konsonantischen Kontrasten, zeigen sich auch bei den Diskriminationsfähigkeiten zwischen verschiedenen nicht-nativen Vokalkategorien Veränderungen während des ersten Lebensjahres. So fanden Polka und Werker (1994) während des ersten Lebensjahres eine Abnahme der Diskriminationsleistungen auch für nicht-muttersprachliche vokalische Kontraste. Im Vergleich zu Konsonantenkontrasten setzt die Abnahme der Diskriminationsfähigkeit für Vokale offensichtlich früher ein. Während vier Monate alte Säuglinge die nicht-nativen Vokalkontraste noch gut unterscheiden konnten, hatte diese Fähigkeit bei 6 bis 8 Monate alten Säuglingen bereits deutlich nachgelassen und war bei 10 bis 12 Monate alten Kindern kaum noch nachzuweisen. Dieser Befund stimmt zeitlich mit dem Auftreten des perzeptuellen Magneteffekts überein, so dass man davon ausgehen kann, dass im Vokalbereich bereits im Alter von 6 Monaten deutliche Einflüsse des Inventars der Umgebungssprache zu beobachten sind.

Auch über die Diskriminationsfähigkeit für einzelne Segmente hinaus findet sich Evidenz, dass sich Säuglinge in der zweiten Hälfte des ersten Lebensjahres weitreichende Kenntnisse über das segmentale Inventar ihrer Muttersprache angeeignet haben. So sind Kinder im Alter von 9 Monaten in der Lage, ihre Muttersprache allein anhand von Unterschieden im segmentalen Inventar von einer Fremdsprache zu unterscheiden, während 6 Monate alte Säuglinge dies offensichtlich noch nicht können (Jusczyk, Friederici, Wessels, Svenkerud & Jusczyk, 1993).

Diese Veränderungen in der Wahrnehmung und das zunehmende Wissen über das segmentale Inventar der Muttersprache, die vor allen Dingen in der zweiten Hälfte des ersten Lebensjahres dokumentiert wurden, sind eine wesentliche Voraussetzung für den weiteren Erwerb phonologischer Regeln, sowohl für den Aufbau phonologischer Regeln als auch für den Aufbau phonologischer Repräsentationen. Um zu Regeln und Repräsentationen zu gelangen, die dem angenommenen System der Zielsprache entsprechen, muss das Kind über die konkreten phonetisch-akustischen Merkmale seines Inputs abstrahieren und Lautklassen bilden, über die diese Regeln und

Repräsentationen operieren. Mit der Schärfung der Wahrnehmungsmechanismen auf muttersprachlich-relevante Informationen verfügt das Kind über die entscheidenden Mechanismen zur Bildung dieser Klassen und zur Zuordnung der Segmente des akustischen Signals zu diesen Klassen.

6.2.2 Der Erwerb zielsprachlichen kombinatorischen und phonotaktischen Wissens

In der zweiten Hälfte ihres ersten Lebensjahres verfügen Kinder neben Wissen über das spezifische Phonem- bzw. Lautinventar ihrer Sprache auch bereits über Kenntnisse von Kombinationsmöglichkeiten dieser Elemente und über die Häufigkeiten, mit denen mögliche Kombinationen auftreten.

In einer ersten Arbeit im Rahmen dieser Fragestellung konnten Jusczyk et al. (1993) zeigen, dass amerikanische Säuglinge im Alter von 9 Monaten in der Lage sind, niederländische Wörter von englischen Wörtern zu unterscheiden, wenn die niederländischen Wörter gegen phonotaktische Regularitäten des Englischen verstoßen, z. B. im Anlaut die Konsonantenverbindung [kn] enthalten. Wichtig ist anzumerken, dass das in den niederländischen Wörtern auftretende segmentale Inventar mit dem englischen kompatibel war, so dass sichergestellt ist, dass die Unterscheidung zwischen den beiden Sprachen nicht allein auf dem Auftreten nicht-nativer Lautsegmente basiert. Da die Säuglinge bei *low-pass* gefilterten Stimuli keine entsprechende Präferenz für die muttersprachlichen Wörter zeigten, ist davon auszugehen, dass auch keine prosodischen Unterschiede zwischen den beiden Sprachen für diese Diskriminationsleistung verantwortlich war. Bei 6 Monate alten Kindern zeigte sich noch keine Differenzierung zwischen den niederländischen und den englischen Stimuli, so dass davon auszugehen ist, dass sie noch nicht auf die phonotaktischen Regeln ihrer Muttersprache reagieren. Dieses Muster lässt darauf schließen, dass sich der Aufbau phonotaktischen Wissens in wesentlichen Zügen zwischen dem sechsten und dem neunten Lebensmonat vollzieht.

Für eine Zunahme der Sensivität für die Häufigkeit von Segmentkombination in diesem Altersbereich sprechen auch Befunde von Jusczyk, Luce und Charles-Luce (1994). Anders als Jusczyk et al. (1993) ging es in dieser Arbeit nicht um phonotaktische Beschränkungen im eigentlichen Sinn, sondern um die Frage, ob und

wann Kinder Wissen über die Häufigkeit bestimmter Segmentabfolgen erworben haben. Sie präsentierten ihren Untersuchungsteilnehmern CVC-Silben, bei denen zum einen jeder einzelne Konsonant und Vokal eine hohe Auftretenshäufigkeit für die jeweilige Position innerhalb der CVC-Struktur hatte und zum anderen die Auftretenshäufigkeiten der Biphone, die in der Onset-Nukleus-Kombination und der Nukleus-Koda-Kombination auftraten, in realen englischen Wörtern relativ hoch war. In der zweiten Untersuchungsbedingung wurden CVC-Silben präsentiert, die in beiden Parametern relativ geringe Werte aufwiesen. Neun Monate alte Säuglinge präferierten in der Untersuchung signifikant die Silben mit hohen Auftretenshäufigkeiten während sich bei 6 Monate alten Säuglingen noch kein Unterschied zwischen den verschiedenen Silbentypen zeigte. In Übereinstimmung mit den oben zitierten Befunden von Saffran und Mitarbeitern (Saffran et al., 1999) zeigen diese Daten, dass 9 Monate alte Säuglinge nicht nur Wissen über extreme Unterschiede in Kookkurrenzmustern (d. h. Auftreten gegenüber Nicht-Auftreten) aufgebaut haben, sondern dass sie darüber hinaus Wissen über relativ subtile Häufigkeiten von Kookkurrenzen erworben haben.

Friederici und Wessels (1993) gingen einen Schritt weiter, indem sie die Sensitivität von Säuglingen für die positionalen Beschränkungen phonotaktischer Regeln untersuchten. Phonotaktische Regeln sind in vielen Sprachen spezifisch für die verschiedenen Silbenpositionen, denn die Beschränkungen für Silbenonset und Silbenkoda können voneinander verschieden sein. So sind in Sprachen, die Konsonantenbündel innerhalb der Silbe zulassen, häufig spiegelverkehrte Beschränkungen für den Silbenonset und die Silbenkoda zu beobachten. Beispielsweise ist im Deutschen die Konsonantenfolge [tr] im Silbenonset zulässig (z.B. *Tränke*), nicht jedoch in der Silbenkoda. Hier tritt demgegenüber die umgekehrte Abfolge [rt] auf (z.B. *Wort*), die ihrerseits wiederum im Silbenonset nicht zulässig ist.

Friederici und Wessels fanden, dass niederländische 9 Monate alte Säuglinge Silben, mit legalen Konsonantenbündeln im Onset bzw. Koda von Silben diskriminieren, in denen dieselben Konsonantenbündel in der Silbenposition auftauchen, in der sie im Niederländischen nicht erlaubt sind (z.B. das Bündel *br* in legaler Position in der Silbe *bref* gegenüber illegaler Position in der Silben *febr*). Bei 6 Monate alten Säuglingen zeigte sich diese Fähigkeit noch nicht.

Der Erwerb phonotaktischen Wissens setzt über die reinen distributionellen Analysemechanismen hinaus auch die Fähigkeit voraus, den Input in Segmente zu zergliedern, innerhalb derer die phonotaktischen Restriktionen wirksam sind. Dies sind wie gesagt Silben, könnten jedoch auch erste einsilbige Wörter sein. Sicherlich verfügen Kinder im Alter von 9 Monaten bereits über die ersten Repräsentationen – zumindest Formrepräsentationen – häufiger Wörter.

6.2.3 Der Aufbau erster Wortform-Repräsentationen

Eine der ersten Phonemsequenzen, die Kinder offensichtlich in ihrem sprachlichen Input als vertraute Elemente wiedererkennen, ist der eigene Name. Mandel, Jusczyk & Pisoni (1995) fanden dass Säuglinge im Alter von 4,5 Monaten länger mehrfachen Wiederholungen des eigenen Namens zuhörten, als fremden Namen, die sowohl in Silbenzahl als auch im Betonungsmuster dem eigenen Namen vergleichbar waren. Diese Befunde weisen darauf hin, dass Kinder bereits im Alter von 4 bis 5 Monaten erste Repräsentationen von Lautsequenzen aus dem Input aufgebaut haben. Nun spielt der eigene Name im Input wahrscheinlich eine herausragende Rolle, zum einen durch eine sicherlich recht hohe Frequenz, zum anderen möglicherweise auch durch eine relativ hohe prosodische Salienz. Säuglinge werden häufig mit dem eigenen Namen angesprochen, um die Aufmerksamkeit des Kindes auf den Sprecher zu lenken, wobei typischerweise die prosodischen Merkmale kindgerichteter Sprache wie Längung, hohe Stimmlage usw. eingesetzt werden. Außerdem wird der Name des Kindes wahrscheinlich relativ häufig isoliert geäußert, was ebenfalls zur akustischen Salienz beiträgt. In dieser Hinsicht, ist der eigene Name wahrscheinlich kein prototypisches sprachliches Element, aber die gefundene sehr frühe Sensitivität für den eigenen Namen zeigt, dass die Charakteristika des Inputs eine klare Relevanz für die Schnelligkeit, mit der eine Repräsentation für eine Lautsequenz aufgebaut wird, besitzen.

Zusammenfassend lässt sich also folgendes festhalten: im letzten Viertel des ersten Lebensjahres verfügt das Kind bereits über umfangreiches Wissen über die nicht-prosodische phonologische Struktur seiner Zielsprache. Es hat Kenntnisse über das native Lautinventar und über die Auftretenshäufigkeit von Phonemkombinationen. Es verfügt über die Voraussetzungen zum Aufbau phonologischer Repräsentationen und

hat wahrscheinlich auch schon die ersten Repräsentationen aufgebaut. Im folgenden soll nun weiter der Frage nachgegangen werden, inwieweit dieses Wissen einen Beitrag leistet zum Aufbau von Wissen auf höheren Ebenen.

6.3 Die Nutzung segmental-distributionellen Wissens für die Verarbeitung des Inputs und den Aufbau höheren sprachlichen Wissens

6.3.1 Die Erkennung von Wortgrenzen

6.3.1.1 Allophonie

Wie im vorherigen Kapitel dargestellt, kann die lautliche Variation innerhalb einer Phonemkategorie in zweierlei Hinsicht für das Erkennen von Wortgrenzen relevant sein. Zum einen können bestimmte allophonische Varianten an bestimmte Positionen innerhalb eines Wortes gebunden sein, so wie im vielzitierten Beispiel der Aspiration von Verschlusslauten im Englischen (z. B. Church, 1987), die ausschließlich silbeninitial vorkommt. Gleichzeitig weisen dann nicht-aspirierte Verschlusslaute auf das Silbenende und damit ein potentielles Wortende hin. Im Deutschen könnte das Merkmal der Stimmhaftigkeit bei Verschlusslauten eine ähnliche Funktion ausüben. Durch die Auslautverhärtung finden sich silbenfinal nie stimmhafte Plosive, das Auftauchen eines solchen markiert also immer einen Silbenanfang und damit auch einen potentiellen Wortanfang. Allerdings ist in diesem Fall der Umkehrschluss nicht zulässig, da stimmlose Konsonanten sowohl silbeninitial als auch silbenfinal auftreten. In diesem Sinne könnten allophonische Varianten und auch phonotaktische Beschränkungen als direkte Markierungen relevanter Grenzen dienen.

Der Frage, inwieweit Säuglinge allophonische Variation als Hinweis auf Wortgrenzen nutzen können, gingen Jusczyk, Hohne und Baumann (1999) nach. Sie nutzen dafür das Stimuluspaar *night rate* vs. *nitrate*. Die wesentliche allophonische Variation zwischen diese Stimuli findet sich in den Realisierungen von /t/ und /r/. In *nitrate* wird das erste /t/ aspiriert und der Verschluss wird gesprengt. Das folgende /r/ ist

stimmlos, was ein Kennzeichen dafür ist, dass es als Teil eines Konsonantenbündels auftritt. Dagegen ist in *night rate* das erste /t/ nicht aspiriert und der Verschluss wird nicht gelöst, wie es im Englisch wortfinal üblich ist. Das folgende /t/ ist stimmhaft, was andeutet, dass es silbeninitial auftritt. Wie Hohne und Jusczyk (1994) zeigten, sind englische Säuglinge im Alter von 2 Monaten bereits fähig, diese allophonen Unterschiede zu diskriminieren. Akustische Messungen der im Experiment verwendeten Stimuli zeigten die oben beschriebenen phonetischen Merkmale.

In der Untersuchung zeigte sich, dass 9 Monate alte Säuglinge diese allophonen Hinweise noch nicht verwenden, um zwischen dem einzelnen Wort *nitate* und der Wortfolge *night rate* zu differenzieren. Im Gegensatz dazu unterschieden Kinder im Alter von 10.5 Monaten zwischen den beiden allophonisch verschiedenen Stimuli. Allerdings sind diese Befunde noch kein direkter Hinweis darauf, dass die allophonen Merkmale auch tatsächlich zur Segmentierung eingesetzt werden. Sie lassen keinen weitergehenden Schluss zu, als dass Kinder diesen Alters in der Lage sind, die positional beschränkten Varianten von Phonemen zu differenzieren. Damit haben sie allerdings eine Voraussetzung dafür, diese positionalen Restriktionen zu erwerben und diese dann für die weitere Sprachverarbeitung einzusetzen.

Allerdings fanden Mattys und Jusczyk (2001b), dass 8 Monate alte Säuglinge nicht reagieren, wenn sie mit CVC-Wörtern familiarisiert worden waren (z.B. *dice*) und in später präsentierten Texten, die gleiche Phonemsequenz über zwei Wörter verteilt vorkam (z.B. *cold ice*). Akustische Messungen von Dauer, Frequenz und Amplitude zeigten, dass es keine massiven Unterschiede zwischen den Realisierungen dieser Phonemsequenz in *cold ice* und *dice* gab, wobei die Sequenz in beiden Fällen innerhalb einer Textpassage ausgesprochen worden war. Daraus schließen die Autoren, dass subtilere allophonische Unterschiede zwischen den beiden Phonemsequenzen verhinderten, dass die Kinder die familiarisierte Einheit in den Textpassagen wiederentdeckten. Der Grund für Widersprüchlichkeit dieser Befunde zu denen von Jusczyk et al. (1999) ist zunächst unklar. Ein möglicher Grund könnte sein, dass das in die Textpassagen eingebettete Wort, das phonemisch dem *Rhyme* des einsilbigen Familiarisierungsworts entsprach, stets mit einem Vokal begann. Möglicherweise weisen wortinitial und wortmedial verwendete Vokale größere akustisch-phonetische

Unterschiede auf als Konsonanten in diesen verschiedenen Wortpositionen. Dies lässt sich aus den Befunden von Nakatani und Dukes (1977) schließen, die fanden, dass erwachsene Probanden Sequenzen mit ambigen Wortgrenzen relativ gut disambiguieren, wenn eins der Worte initial einen Vokal aufweist. Wie phonetische Messungen hier zeigten, ging initialen Vokalen sehr zuverlässig ein glottaler Verschlusslaut voraus. Da Mattys und Jusczyk offensichtlich segmentale Unterschiede zwischen den Sequenzen nicht analysiert haben, könnte dies eine Quelle der Unterscheidungsmöglichkeit für die Säuglinge gewesen sein. Zum zweiten ist möglicherweise auch das Phänomen der Koartikulation dafür verantwortlich, dass die Kinder nicht auf die familiarisierte Sequenz reagierten, wenn diese in den Textpassagen durch eine Wortgrenze geteilt wurde. Wahrscheinlich gibt es in der Sequenz *dice* deutliche Koartikulationsmerkmale bei der Realisierung des initialen Verschlusslautes und des folgenden Vokals. Diese fehlen wahrscheinlich, wenn zwischen den beiden Lauten eine Wortgrenze liegt, so dass es phonetische Unterschiede zwischen der familiarisierten Form und der im Text vorkommenden gibt.

Ob Koartikulation von Säuglingen in dieser Hinsicht perzeptuell ausgenutzt wird, wird in einer Untersuchung von Johnson und Jusczyk (2001) thematisiert. In dem analog zu dem von Saffran, Aslin und Newport (1996) aufgebauten Experiment wurden die Kinder zunächst mit einem kontinuierlichen Sprachstrom familiarisiert, der in variabler Abfolge Dreisilbler enthielt. In der Testphase wurden den Kindern Zweisilbler präsentiert, die entweder aus zwei aufeinanderfolgenden Silben der Dreisilbler konstruiert worden waren (Wörter) oder aus den End- und Anfangsilben zweier verschiedener Dreisilbler (Teilwörter). Im kontinuierlichen Sprachstrom waren die Silben der Wörter immer adjazent präsentiert worden, während dies bei den Teilwörtern nur in 30 % der Präsentationen der Fall war. Um die Wirksamkeit koartikulatorischer Hinweise für die Segmentierung zu überprüfen, wurden die Teilwörter zusammenhängend gesprochen, während alle anderen Sequenzen aus isoliert gesprochenen Silben zusammengeschnitten wurden. Mit dieser Vorgehensweise wurde ein Konflikt zwischen den rein distributionellen Hinweisen erzeugt, nach denen ausschließlich die Dreisilbler eine Einheit darstellen sollten, und den koartikulatorischen Merkmalen, nach denen trotz der geringeren Übergangswahrscheinlichkeit die Teilwörter eine Einheit darstellen.

In den Ergebnissen von 8 Monate alten Kindern zeigte sich eine Präferenz für die Wörter, also diejenigen Einheiten, die nach den rein distributionellen Hinweisen als Einheit gelten. Da sich sowohl in Saffran et al. (1996) als auch in einer Replikation dieser Studie innerhalb der Untersuchungsreihe von Johnson und Jusczyk (2001) stets eine Präferenz für die Teilwörter, also ein Neuheitseffekt gezeigt hatte, lassen sich die Daten dahingehend interpretieren, dass die Kinder in diesem Experiment die koartikulierten Silbensequenzen als bekannte Stimuli behandelten, die ausschließlich nach distributionellen Kriterien zusammengehörigen jedoch als unbekannte oder zumindest unbekanntere. Dieses Muster deutet darauf hin, dass die Koartikulationsmerkmale einen stärkeren Effekt auf die Segmentierung des Sprachstroms haben als die distributionellen Merkmale.

6.3.1.2 Statistisches Wissen über Lautabfolgen und phonotaktisches Wissen

Der Einsatz phonotaktischen Wissens wurde insbesondere im Zusammenhang mit der Frage nach der Segmentierung des sprachlichen Inputs diskutiert. Mattys, Jusczyk, Luce und Morgan (1999) untersuchten, ob 9 Monate alte Säuglinge, die Häufigkeit, mit der bestimmte Konsonantenabfolgen wortintern bzw. über Wortgrenzen hinweg auftreten, zur Wortsegmentierung nutzen können. Dazu präsentierten sie den Kindern zweisilbige CVC-CVC Pseudowörter. Variiert wurde in den Pseudowörtern, ob die interne Konsonantenfolge aus Konsonanten bestand, die nach einer Zählung eines kindgerichteten englischen Sprachkorpus häufiger wortintern als wortübergreifend vorkommen, oder ob diese Folge aus Konsonanten bestand, die in dem Korpus häufiger wortübergreifend als wortintern vorkamen. Gleichzeitig wurde das metrische Muster der Zweisilbler variiert, die entweder auf der ersten oder auf der zweiten Silbe betont sein konnten.

Bei den trochäischen Pseudowörtern präferierten die Kinder die Sequenzen mit häufiger wortintern auftretenden Konsonantenabfolgen gegenüber den Sequenzen, die häufiger wortübergreifend auftretende Konsonantenabfolgen enthielten. Bei den jambischen Pseudowörtern zeigte sich dagegen die entgegengesetzte Präferenz mit längeren Orientierungszeiten für die Sequenzen mit wortübergreifenden Konsonantenabfolgen. Dieses Muster zeigt deutlich, dass die Kinder sowohl die

prosodische als auch die phonotaktische Information nutzen. Sie bevorzugten jeweils die Stimuli, in denen beide Informationstypen kongruent sind. Nach den in Kapitel 4 berichtete Befunden sollten trochäische Stimuli von Kindern dieses Alters bereits als Einheit wahrgenommen werden, entsprechend ist hier die häufiger wortintern vorkommende Konsonantenabfolge konform mit der prosodischen Struktur. Jambische Stimuli sollten dagegen bei Kindern dieses Alters vor der starken Silbe segmentiert werden, in diesem Fall stützt die häufiger wortübergreifend vorkommende Konsonantenabfolge am potentiellen Segmentierungspunkt das Ergebnis der prosodischen Analyse.

Darüber hinaus konnten Mattys und Jusczyk (2001a) zeigen, dass Kinder die Übergangswahrscheinlichkeiten zwischen Konsonanten auch nutzen, um Wörter aus kontinuierlicher Sprache zu segmentieren. In dieser Untersuchung präsentierten die Forscher ebenfalls 9 Monate alten Kindern CVC-Silben, die so in Textpassagen eingebettet worden waren, dass der kritischen CVC-Silbe jeweils ein Wort voranging, das auf einen Konsonanten auslautete und ein Wort folgte, das mit einem Konsonanten begann. Variiert wurde dabei, ob der der CVC-Silbe vorangehende bzw. folgende Konsonant, mit dem Silbenonset bzw. der Silbencoda eine häufiger wortintern oder eine häufiger wortübergreifend vorkommende Konsonantenabfolge darstellte. Es wurden dabei die gleichen Kombinationen verwendet wie in Mattys et al. (1999). Jedem Kind wurde zunächst eine Textpassage mit einer CVC-Silbe mit wortinternem Kontext und eine Textpassage mit einer CVC-Silben mit wortübergreifendem Kontext präsentiert. In der eigentlichen Testphase wurden die isolierten CVC-Silben sowie einige in den Textpassagen nicht vorkommenden Ablenker präsentiert. Es zeigten sich längere Orientierungszeiten für die Silben, die im wortübergreifenden Kontext präsentiert worden waren, also in dem Kontext, der hypothesengemäß die stärkeren Segmentierungshinweise enthielt.

Diese Befunde liefern deutliche Evidenz dafür, dass 9 Monate alte Säuglinge in der Lage sind, ihr bereits vorhandenes Wissen über Regularitäten von Segmentfolgen für die Extraktion von wortähnlichen Einheiten aus dem Input heranzuziehen. Um dieses Wissen zu erwerben, muss das Kind bereits über erste Wortsegmentierungsstrategien verfügen. In dieser Hinsicht können die hier untersuchten Fähigkeiten nur als eine Ergänzung bereits existenter Strategien angesehen

werden. Als eigentlicher erster Einstieg in die Lösung des Segmentierungsproblems sind sie ungeeignet, sie könnten aber durchaus dazu beitragen, von den ersten Segmentierungsstrategien nicht zu bewältigende Problemfälle zu lösen.

6.3.1.3 Kookkurrenzen von Silben

In Abschnitt 6.1.3 wurden Befunde berichtet, die zeigen, dass Säuglinge extrem sensitiv für die Übergangswahrscheinlichkeiten zwischen Silben sind. Es stellt sich nun die Frage, ob und wenn ja ab welchem Alter diese Sensitivität, dazu führt, dass Abfolgen mit hoher Übergangswahrscheinlichkeit tatsächlich als Einheit wahrgenommen werden.

Deutliche Evidenz für die Rolle von Übergangswahrscheinlichkeiten auf die Erkennung von Einheiten liefern die Befunde von Morgan und Saffran (1995). In ihrer Untersuchung mit der *Noise-Detection Technik* stellten sie fest, dass 9 Monate alte Kinder nur dann langsamer auf ein Störgeräusch innerhalb eines Silbenpaares reagierten, wenn ihnen dieses Silbenpaar zuvor mehrfach sowohl in derselben Abfolge als auch im selben rhythmischen Muster präsentiert worden war. Demgegenüber reagierten 6 Monate alte Säuglinge noch nicht auf die sequentielle sondern lediglich auf die rhythmische Information.

Dass auch in längere variable Kontexte eingebettete konstante Silbenabfolgen als Einheit extrahiert werden können, zeigt ein weiterer Befund aus der Studie von Jusczyk et al. (1999). In dieser Serie von Experimenten stellte sich unter anderem heraus, dass 7,5 Monate alte Säuglinge eine Silbe, die zunächst in einer Textpassage als starke Silbe eines zweisilbigen Wortes präsentiert worden war, bei einer späteren isolierten Darbietung wiedererkennen, wenn die Silbe im Text Teil eines jambischen Wortes war, nicht aber wenn sie Teil eines trochäischen Wortes war. Zudem erkannten gleichaltrige Kinder bei einer Familiarisierung mit einer einzelnen Silbe diese in einer Textpassage wieder, wenn sie dort die starke Silbe eines jambischen Wortes war, nicht aber wenn sie die starke Silbe eines trochäischen Wortes war. Die Autoren vermuteten, dass dieser Unterschied mit den verschiedenen Distributionen der starken Silbe in den trochäischen und den jambischen Wörtern zu hat: in den trochäischen Wörtern folgte der starken Silbe immer dieselbe Silbe, nämlich die zweite Wortsilbe. In den jambischen Wörtern dagegen war die der starken Silbe folgende Silbe nicht mehr Teil

des jambischen Wortes und somit in Abhängigkeit vom folgenden Wort variabel. Diese Variabilität des Kontextes könnte dazu beigetragen haben, dass die starke Silbe des jambischen Wortes aus dem Kontext herausgelöst wird und als einsilbige Einheit repräsentiert wird.

Diese Vermutung wurde mit zwei weiteren Experimenten überprüft, in denen Textpassagen benutzt wurden, in denen dem kritischen jambischen Wort immer dasselbe unbetonte Funktionswort folgte (z.B. *guitar is*). Damit unterlag die starke Silbe des jambischen Wortes den gleichen Distributionsbedingungen wie die starke Silbe des trochäischen Wortes: es folgte ihr im Text stets die gleiche unbetonte Silbe. Im ersten dieser Experimente wurde 7,5 Monate alten Säuglingen wiederum in der Familiarisierungsphase die isolierte starke Silbe aus den jambischen Wörtern geboten und in der Testphase die veränderten Textpassagen. In diesem Fall zeigten die Kinder – wie bei den starken Silben der trochäischen Wörter – keine Anzeichen, dass sie die starke Silbe in den Textpassagen wiedererkannten. Noch deutlichere Hinweise, dass die Kinder die konstanten Silbenpaare in den Textpassagen als Einheiten extrahierten und repräsentierten, ergab sich im zweiten Folgeexperiment. Hier wurden den Kindern zuerst die veränderten Textpassagen präsentiert und in der Testphase Pseudowörter, die aus der starken Silbe des jambischen Wortes und der folgenden schwachen Silbe gebildet worden waren (*taris*). Es zeigte sich ein Familiarisierungseffekt für diese Pseudowörter gegenüber den Zweisilblern, die in den Textpassagen nicht vorgekommen waren.

Zusammengenommen zeigen diese Befunde deutlich, dass Kinder zwischen dem siebten und dem neunten Lebensmonat Analyse- und Strukturierungsmechanismen entwickelt haben, die es ihnen erlauben, anhand sequentieller Information Einheiten zu bilden und zwar auch dann, wenn diese Sequenzen in größeren Kontexten vorkommen. Gerade die zuletzt genannten Befunde von Jusczyk et al. (1999) weisen erneut auf eine enorme Leistungsfähigkeit dieser Mechanismen hin, denn die kritische Silbenabfolge wurde in jeder Textpassage nur sechsmal präsentiert. Bei einer Länge der einzelnen Textpassagen von durchschnittlich 20 Sekunden und einer Familiarisierungsdauer von 45 Sekunden pro Textpassage bedeutet dies, dass die Kinder die Silbenabfolge ca. 13 mal gehört hatten.

6.3.1.4 Erkennung und Nutzung hochfrequenter Phonemfolgen

Die vorherigen Abschnitte haben gezeigt, dass Säuglinge in der zweiten Hälfte ihres ersten Lebensjahres über Verarbeitungs- und Analysemechanismen verfügen, die es ihnen erlauben, in ihrem sprachlichen Input häufig wiederkehrende Muster zu erfassen und diese Information für die Strukturierung des Inputs zu nutzen. Diese Fähigkeit zeigt sich nicht nur bei konzentrierter Präsentation sich wiederholender Sequenzen unter experimentellen Bedingungen, sondern prägt offensichtlich auch die Verarbeitung der alltäglichen sprachlichen Umgebung des Säuglings. Wie sonst wäre es zu erklären, dass bereits 9 Monate alte Kinder Wissen über die Häufigkeit von Phonemkombinationen in ihrer Sprache haben und phonotaktisch legale von phonotaktisch nicht zulässigen Konsonantenbündeln unterscheiden können.

In diesem Abschnitt soll nun weiter dafür argumentiert werden, dass diese distributionellen Fähigkeiten dazu führen sollten, dass Säuglinge bereits sehr früh mentale Repräsentationen für solche Wortformen aufbauen, die in ihrem Input hochfrequent sind. Diese ersten mentalen Repräsentationen können von den Kindern wiederum in ihrem sprachlichen Input wiedererkannt werden und unterstützen in dieser Weise als eine Art Ankerpunkt (Valian & Coulson, 1988) weitere Analyseprozesse, in dem sie beispielsweise Informationen über die Grenzen sie umgebender Wörter liefern. Dass die Input-Frequenz für den Aufbau lexikalischer Repräsentationen eine Rolle spielen könnte, wurde bereits im Zusammenhang mit den Befunden diskutiert, nach denen Säuglinge bereits im Alter von viereinhalb Monaten in der Lage sind, ihren eigenen Namen zu erkennen (Mandel et al., 1995). Wenn eine hohe Input-Frequenz den Aufbau mentaler Repräsentationen stützt ist zudem zu erwarten, dass auch Funktionswörter zu den ersten mentalen Repräsentation des Kindes gehören, da diese im allgemeinen zu den häufigsten Wortformen der Sprache gehören. So berichten Morgan, Shi & Allopenna (1996), dass im Kucera-Francis Korpus, einem gängigen Korpus englischer schriftsprachlicher Texte, die 50 häufigsten Wörter der Klasse der Funktionswörter zugehören und insgesamt 40% aller Token dieses Korpus Funktionswörter sind.

Die Annahme früher mentaler Repräsentationen gerade für Funktionswörter scheint zunächst unplausibel, wenn man die Sprachproduktion betrachtet. Die ersten Mehrwortäußerungen von Kindern sind gerade dadurch gekennzeichnet, dass sie im

wesentlichen aus Inhaltswörtern bestehen bei fehlenden Funktionswörtern und Flexionsendungen. Diese Beobachtung legt die Vermutung nahe, dass Funktionswörter später erworben werden als Inhaltswörter, was von einigen Autoren wiederum mit ihrer geringen phonologischen Prominenz im Input, vor allen Dingen der Tatsache, dass sie im allgemeinen unbetont sind, und darauf basierenden Wahrnehmungsschwierigkeiten in Verbindung gebracht wurde (s. z.B. Bates & Goodman, 1999; Gleitman & Wanner, 1982).

Neben ihrer hohen Frequenz weisen Funktionswörter jedoch noch weitere Eigenschaften auf, die sie möglicherweise gerade den Verarbeitungsmöglichkeiten jüngerer Kinder gut zugänglich machen. Funktionswörter sind phonologisch meist relativ einfach aufgebaut, d. h. sie sind häufig einsilbig und weisen eine einfache Silbenstruktur auf. Zudem treten Funktionswörter sehr häufig an den Rändern prosodischer Einheiten, wie Intonationsphrasen und phonologischen Phrasen auf. Durch die relativ deutliche prosodische Markierung dieser Grenzen sind Elemente, die den Grenzen folgen oder vorangehen möglicherweise einfacher zu extrahieren als Elemente, die innerhalb von prosodischen Einheiten auftreten (s. Christophe et al. 1997; Gerken, 2001).

Inzwischen liegt denn auch eine Vielzahl von Befunden vor, die der Annahme besonderer Schwierigkeiten bei der Perzeption und der Verarbeitung von Funktionswörtern und anderen unbetonten Elementen bei Säuglingen widersprechen. So fanden Jusczyk und Thompson (1978), dass zwei Monate alte Säuglinge Lautkontraste in unbetonten Silben genauso gut diskriminieren können wie Lautkontraste in betonten Silben. Unsere eigene, in Abschnitt 6.1.2.1 bereits näher beschriebene Untersuchung weist in die gleiche Richtung. Im gleichen Zeitraum, in dem Säuglinge in der Lage sind, Nomen in kontinuierlicher Sprache zu entdecken, haben sie diese Fähigkeit auch für unbetonte Elemente der geschlossenen Klasse. Zudem häuft sich die Evidenz dafür, dass aus dem Fehlen von grammatischen Morphemen in ersten Mehrwortäußerungen nicht direkt geschlossen werden kann, dass den Kindern das notwendige sprachliche Wissen zu einer korrekten Verwendung dieser Elemente fehlt. Gerken und Kollegen fanden in diversen Studien Unterschiede in den Reaktionen von Kindern, die sich noch in der Einwortphase befanden, auf Sätze, die korrekt gebrauchte Funktionsmorpheme enthielten, gegenüber Sätzen in denen diese fehlten oder durch im Kontext nicht

korrekte Morpheme bzw. sinnlose Silben ersetzt worden waren (Gerken, Landau & Remez, 1990; Gerken & McIntosh, 1993; Shafer, Shucard, Shucard & Gerken, 1998; Shady & Gerken, 1999; Shipley, Smith, and Gleitman, 1969). Auch Befunde von Saffran (2001), die bereits bei 8 Monate alten Säuglingen unterschiedliche Reaktionen auf kurze Satzrahmen mit realen englischen Wörtern gegenüber solchen, in denen alle Wörter durch Unsinn-Silben ersetzt worden waren, fand, deuten auf eine frühe Sensitivität für Funktionswörter hin, denn die Satzrahmen bestanden zum Großteil aus Funktionswörtern. Santelmann und Jusczyk (1998) berichten, dass bereits 18 Monate alte englische Säuglinge Sätze präferieren, in denen Flexionsendungen korrekt gebraucht werden, gegenüber Sätzen mit fehlenden Flexionsendungen (s. a. Golinkoff, Hirsh-Pasek & Schweisguth, 2001 für analoge Befunde bei Derivationsuffixen). Weissenborn, Höhle, Kiefer und Cavar (1998) konnten zeigen, dass deutschsprachige Kinder im Alter von 2,5 Jahren bereits über Wissen über den Zusammenhang des Auftretens eines Komplementierers und der Stellung des finiten Verbs im Komplementsatz haben, lange bevor Konstruktionen dieses Typs von den Kindern spontan verwendet werden.

Alle diese Befunde zeigen, dass bereits früh während des Spracherwerbs grammatische Morpheme wahrgenommen und verarbeitet werden und dass Kinder bereits in der zweiten Hälfte des zweiten Lebensjahres weitreichendes Wissen über die Verwendungsbeschränkungen grammatischer Morpheme erworben haben. Verschiedene Autoren haben darüber hinaus die Vermutung geäußert, dass eine Erkennung von Funktionswörtern als vertraute Lautsequenz innerhalb des sprachlichen Inputs durch eine Grenzziehung rechts und links dieser Sequenzen für die Segmentierung einen wichtigen Beitrag liefern können (z.B. Christophe et al., 1997; Gerken, 2001; Morgan et al., 1987; Shady & Gerken, 1999; Shafer et al., 1998).

6.3.1.5 Nutzung von Funktionswörtern als Segmentierungshilfe

Im folgenden soll über eine Untersuchung berichtet werden, die die Hypothese, dass bereits während der vorsprachlichen Phase Funktionswörter erkannt und verarbeitet werden und diese zur Segmentierung des sprachlichen Inputs herangezogen werden, überprüft. In diesem Experiment wurden Kinder entweder mit einer Artikel-Nomen-

Sequenz (*der Kahn*) oder einem jambischen Nomen, das die homophone starke Silbe enthielt wie die Nominalphrase (*Vulkan*). Anschließend wurden die Kinder mit Textpassagen getestet, die nur die starke Silbe der Familiarisierungssequenzen enthielt (*Kahn*). Falls die Kinder die Nominalphrasen durch das Erkennen der Artikelformen als zwei Einheiten repräsentieren, sollten sie die starke Silbe in den Texten wiedererkennen. Demgegenüber sollten sie die jambischen Nomen als Einheit repräsentieren, da die erste Silbe keinem lexikalischen Wort des Deutschen entspricht. Wenn die Kinder dagegen eine reine rhythmische Segmentierungsstrategie anwenden, sollten sie beide Familiarisierungssequenzen vor der starken Silbe segmentieren und daher unabhängig von der Art der Familiarisierung einen Wiedererkennungseffekt für die starke Silbe in den Textpassagen zeigen. Da nach den Befunden Jusczyks et al. (1999) englische Kinder jambische Wörter erst im Alter von ca. 11 Monaten als Einheiten repräsentieren, wurden Kinder dieses Alters in das Experiment einbezogen.

6.3.1.5.1 Experiment 6: Erkennung und Segmentierung von Artikeln bei 11-monatigen Kindern

Probanden

In die Untersuchung wurden die Daten von 48 Kindern, jeweils 24 pro Familiarisierungsbedingung einbezogen. Um diese Gruppenstärke zu erreichen mussten 76 Kinder untersucht werden. Bei der Mehrzahl der ausgeschlossenen Kinder konnte das Experiment nicht beendet werden, der zweithäufigste Ausschlussgrund waren zu kurze Orientierungszeiten. Bei einigen Kindern gab es technische Probleme mit der Durchführung des Experiments. Alle untersuchten Kinder wuchsen in einem monolingualen deutschsprachigen Elternhaus auf und wiesen keine bekannten Hörbeeinträchtigungen auf.

Das durchschnittliche Alter der Kinder der Nomen-Familiarisierung lag bei 11 Monaten und 25 Tagen (Spannbreite: 10 Monate 19 Tage bis 12 Monate 11 Tage), das der Kinder der Phrasen-Familiarisierung bei 11 Monaten und 15 Tagen (Spannbreite: 10 Monate 15 Tage bis 12 Monate 16 Tage). In der Nomenfamiliarisierungs-Gruppe waren 12 Mädchen und 12 Jungen. Die Gruppe der Phrasen-Familiarisierung umfasste 8 Mädchen und 16 Jungen.

Untersuchungsmaterial

Als Untersuchungsmaterial für diese Studie wurden zunächst vier Paare von Nomen des Deutschen ausgewählt. Jedes Paar bestand aus einem zweisilbigen jambischen Nomen und einem einsilbigen Nomen, dessen phonologische Form der starken Silbe des jambischen Nomens entsprach (*Vulkan - Kahn, Hormon - Mohn, Pastor - Tor, Konzil - Ziel*). Zudem wurde für jedes der einsilbigen Nomen eine aus sechs Sätzen bestehende kurze Textpassage konstruiert, in der das kritische Wort in jedem Satz genau einmal vorkam. Dabei wurde darauf geachtet, dass in keiner der Textpassagen der in der Familiarisierungssequenz verwendete definite Artikel des einsilbigen Nomens vorkam.

Das Untersuchungsmaterial wurde von einer Sprecherin des Deutschen eingesprochen. Sie sprach die zweisilbigen Nomen sowie die einsilbigen Nomen zusammen mit dem zugehörigen definiten Artikel (*der/das*) mehrfach isoliert ein. Außerdem wurden die vier Textpassagen von ihr eingesprochen. Die Sprecherin erhielt die Anweisung, die sprachlichen Reize in einer kindgerichteten Sprechweise mit lebhafter Intonation zu produzieren.

Das eingesprochene Material wurde anschließend digitalisiert und mit Hilfe eines Spracheditierungsprogramms wurden die Stimulussequenzen für das Experiment hergestellt. Für die Nomenfamiliarisierung wurden Sequenzen von ca. 30 Sekunden hergestellt, in denen 26 Token verschiedener Realisierungen eines der zweisilbigen Nomen mit einer Pause von 600 ms wiederholt wurden. Für die Phrasenfamiliarisierung wurden in entsprechender Weise ebenfalls ca. 30 Sekunden dauernde Sequenzen hergestellt, in denen Token einer Nominalphrase mit 600 ms Pause zwischen den einzelnen Token wiederholt wurden. Die Textpassagen wurde jeweils als eine Datei abgespeichert. Sie wiesen eine Dauer von ca. 19 Sekunden auf (*Kahn*-Text: 19,53 s; *Tor*-Text: 19,76 s; *Mohn*-Text: 19,71 s; *Ziel*-Text: 19,26 s).

Durchführung

Das Experiment wurde mit dem *Headturn Preference* Paradigma mit Familiarisierungsphase durchgeführt. Für die Familiarisierung gab es zwei verschiedene Experimentalbedingungen. Entweder wurden den Kindern zwei der vier jambischen Nomen wiederholt präsentiert oder es wurden ihnen zwei der vier einsilbigen Nomen in

Verbindung mit dem zugehörigen definiten Artikel präsentiert (*der Kahn, der Mohn, das Tor, das Ziel*). Bei der Phrasenfamiliarisierung wurden nur Phrasen verschiedener Genera gemeinsam präsentiert, um zu verhindern, dass allein das häufige Auftreten der gleichen Phonemsequenz innerhalb der schwachen Silbe der Sequenz die Kinder dazu veranlassen würde, diese aus der Gesamtsequenz herauszulösen. Die Familiarisierung wurde beendet, wenn das Kind für jedes der beiden Nomen bzw. jede der beiden Artikel-Nomen Kombination mindestens eine Orientierungszeit von 30 s erreicht hatte. Direkt im Anschluss daran wurde die eigentliche Testphase gestartet, in der alle Kinder alle vier Textpassagen mit den einsilbigen Nomen jeweils viermal hörten.

Ergebnisse und Diskussion

Wiederum wurden für jedes Kind die Orientierungszeiten für jede Textpassage ermittelt. Daraus wurde pro Kind ein Mittelwert für die Textpassagen mit familiarisierter Silbe und ein Mittelwert für die Textpassagen ohne familiarisierte Silbe berechnet. Bei den Kindern, die mit den Phrasen familiarisiert worden waren, zeigten sich mit durchschnittlich 7,48 s signifikant längere Orientierungszeiten für die Textpassagen, die die familiarisierte Silben enthielten gegenüber den Textpassagen, die die familiarisierten Silben nicht enthielten (durchschnittlich 6,44 s). Dieser Unterschied war statistisch signifikant ($t(df = 23) = 2,55; p < 0,02$). Von den 24 getesteten Kinder zeigten 19 diese Präferenz. Demgegenüber zeigten sich bei den Kindern der Nomen-Familiarisierung mit durchschnittlichen Orientierungszeiten von 7,74 s für die Textpassagen, die familiarisierte Silben enthielten und von 7,33 s für die Textpassagen, die keine der familiarisierten Silben enthielten, kein signifikanter Unterschied ($t(df = 23) = 0,95; p = 0,35$). In dieser Gruppe zeigten 13 der 24 Kinder längeren Orientierungszeiten zu den Textpassagen mit der familiarisierten Silbe.

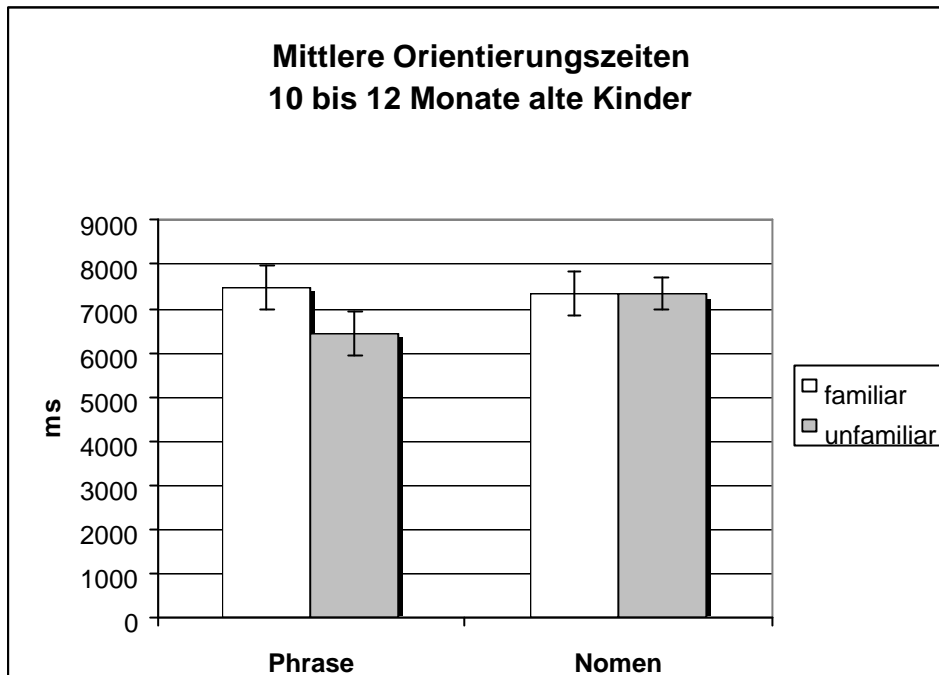


Abbildung 6: Mittlere Orientierungszeiten mit für Texte mit und Texte ohne familiarisierte starke Silben aus Experiment 6

Das Ergebnismuster der älteren Kinder entspricht genau unseren Erwartungen: sie reagieren nur dann auf das einsilbige Nomen im Text, wenn sie die entsprechende Silbe zuvor in Kombination mit einer Artikelform gehört haben, nicht aber wenn diese Silbe Bestandteil eines jambischen Wortes war. Aus diesem Befundmuster lassen sich eine Reihe von Schlüssen ziehen. Erstens, und das ist im Zusammenhang mit unserer Ausgangsfrage der wichtigste Befund, behandeln die Kinder zweisilbige jambische Sequenzen, deren erste Silbe ein eigenständiges sehr gebrauchshäufiges Wort ist - in diesem Fall eine Form des definiten Artikels - anders als jambische Sequenzen, deren erste Silbe kein Wort ist. Die Tatsache, dass die Kinder die starke Silbe der Phrase in den Textpassagen wiedererkannt haben, spricht nach obiger Logik dafür, dass die Phrasen nicht als Einheit repräsentiert wurden, sondern dass die Phrasen in zwei unabhängige Einheiten segmentiert wurden. Man kann also davon ausgehen, dass die Kinder die Artikelformen bereits als vertraute eigenständige Einheiten erkannt haben und damit der Rest der Sequenz ebenfalls eine eigenständige Einheit darstellen musste. Zweitens verfolgen auch deutschsprachige Säuglinge im Alter zwischen 10 und 12

Monaten keine rhythmische Segmentierungsstrategie mehr: sie segmentieren jambische Sequenzen nicht nach rein rhythmischen Gesichtspunkten vor der starken Silbe.

In diesem Zusammenhang ist zu fragen, ob deutsche Kinder sich in dieser Beziehung überhaupt wie ihre englischsprachigen Altersgenossen verhalten. Wir haben in Kapitel 5 zwar gesehen, dass sie ebenso wie englische Kinder trochäische Zweisilbler mit ca. 8 Monaten in Textpassagen wiedererkennen können, über die Verarbeitung jambischer Zweisilbler ist bei deutschen Kindern jedoch bislang nichts bekannt. Aus diesem Grund wurde das oben beschriebene Experiment an einer weiteren Gruppe jüngerer Kinder wiederholt.

6.3.1.5.2 Experiment 7: Segmentierung von Phrasen und jambischen Nomen bei 9 Monate alten Kindern

Probanden

Auch in dieses Experiment wurden die Daten von insgesamt 48 Kindern einbezogen, wiederum 24 pro Familiarisierungsbedingung. Um auf diese Zahl zu kommen war die Untersuchung von insgesamt 67 Kindern nötig. Der Ausschluss der 19 Kinder erfolgte aus den gleichen Gründen wie im vorher beschriebenen Experiment.

Die Kinder in der Nomen-Familiarisierungsgruppe hatten ein durchschnittliches Alter von 9 Monaten und 19 Tagen (Spannbreite: 9 Monate 1 Tag bis 10 Monate 9 Tage). Die Kinder in der Phrasen-Familiarisierung wiesen ein durchschnittliches Alter von 9 Monaten und 21 Tagen auf (Spannbreite: 8 Monate 14 Tage bis 10 Monate 14 Tage). In der Nomenfamiliarisierung befanden sich 9 Mädchen und 15 Jungen. In der Phrasenfamiliarisierung waren 16 Mädchen und 8 Jungen.

Untersuchungsmaterial und Durchführung

Das Stimulusmaterial für diese Untersuchung sowie die Durchführung entspricht dem Experiment bei den 11 Monate alten Kindern.

Ergebnisse und Diskussion

Auch in diesem Fall wurden für jedes Kind die Orientierungszeiten für jede Textpassage ermittelt und daraus die Mittelwerte für die Textpassagen mit familiarisierter starker

Silbe und die Textpassagen ohne familiarisierte Silbe ermittelt. Die Ergebnisse der jüngeren Altersgruppe sehen vollkommen anders aus. Hier zeigte sich bei den mit den Phrasen familiarisierten Kindern kein signifikanter Unterschied zwischen Textpassagen mit familiarisierter Silbe (durchschnittlich 6,99 s) und den Texten ohne familiarisierte Silbe (durchschnittlich 7,52 s) ($t(df = 23) = 0,95, p = 0,34$). Dreizehn der 24 Kinder präferierten dabei die Textpassagen mit familiarisierter Silbe. Demgegenüber zeigten die Kinder, die mit den zweisilbigen Nomen familiarisiert waren, eine hochsignifikante Präferenz für die Textpassagen mit der familiarisierten Silbe (durchschnittlich 8,32 s) gegenüber den Orientierungszeiten für die Textpassagen ohne familiarisierte Silbe (6,72 s; $t(df = 23) = 3,84, p < 0,001$). Bei 20 der 24 Kinder wurde diese Präferenz gefunden.

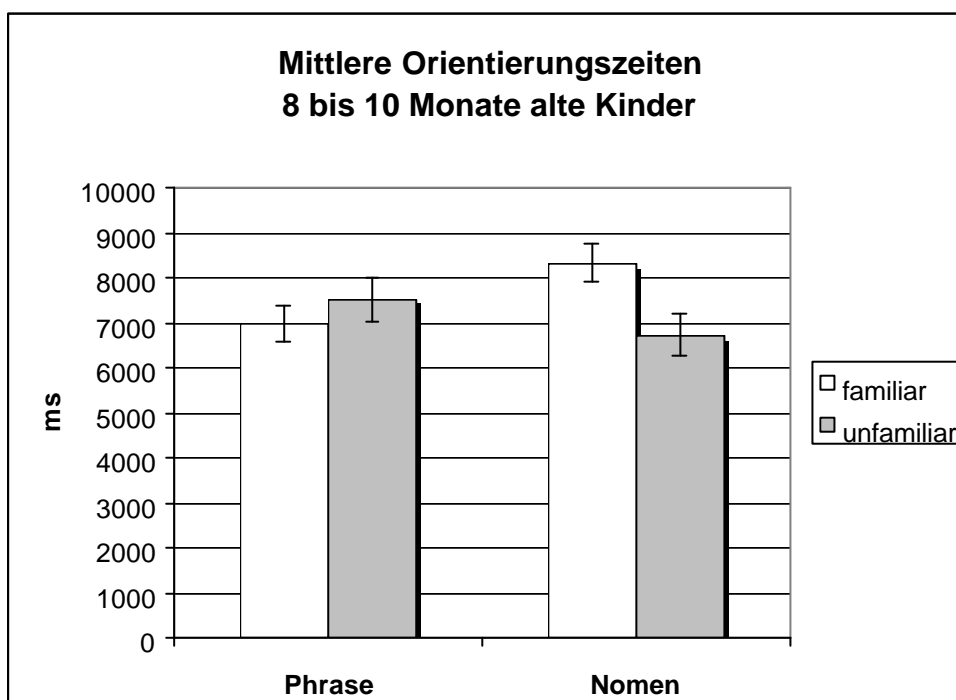


Abbildung 7: Mittlere Orientierungszeiten mit für Texte mit und Texte ohne familiarisierte starke Silben aus Experiment 7

Die Befunde der jüngeren Altersgruppe sind weniger eindeutig zu interpretieren wie dies für die älteren Kinder der Fall war. Zum einen entsprechen die Ergebnisse der Gruppe von Kindern, die mit den zweisilbigen Nomen familiarisiert worden waren genau den Voraussagen der rhythmischen Segmentierungshypothese: die Kinder

reagieren auf die isolierte starke Silbe in den Textpassagen. Dies spricht dafür, dass sie die zweisilbigen jambischen Nomen, die sie während der Familiarisierung gehört haben, am Beginn der starken Silbe segmentiert und als zwei Einheiten repräsentiert haben. Die Ergebnisse der Gruppe, die mit den Phrasen familiarisiert worden sind, spricht allerdings gegen die alleinige Anwendung einer rhythmischen Segmentierungsstrategie. In diesem Fall hätte auch diese Gruppe einen Wiedererkennungseffekt für die starke Silbe in den Textpassagen zeigen sollen. Dieser Effekt fand sich jedoch nicht.

Eine mögliche Erklärung für die unterschiedlichen Ergebnisse der Gruppen der beiden Familiarisierungsbedingungen könnte sein, dass unsere Grundannahme, nach der die zweisilbigen jambischen Nomen und die Artikel-Nomen-Phrasen in Hinblick auf ihre metrischen Eigenschaften identisch sind, falsch war. Um dies zu klären wurde eine akustisch-phonetische Analyse der prosodischen Merkmale der Familiarisierungsstimuli durchgeführt. Für jedes Token der jambischen Nomen und für jedes Token der Artikel-Nomen Phrasen wurde dabei die Länge und die maximale Grundfrequenz der Vokale der beiden Silben bestimmt.

Sowohl für die jambischen Nomen als auch für die Phrasen zeigte sich in der zweiten Silbe eine signifikant längere Vokaldauer als in der ersten Silbe. In den Nomen wies die erste Silbe eine durchschnittliche Vokaldauer von 108 ms auf gegenüber einer durchschnittlichen Vokaldauer von 232 ms in der zweiten Silbe. Der Vokal der ersten Silbe der Phrasen war durchschnittlich 71 ms lang, der Vokal der zweiten Silbe 151 ms. Die Werte für die maximale Grundfrequenz weisen in die gleiche Richtung. In den Nomen fand sich im Vokal der ersten Silbe eine durchschnittliche maximale Grundfrequenz von 332 Hz und in dem der zweiten Silbe von 425 Hz. Innerhalb der Phrasen lag die durchschnittliche maximale Grundfrequenz des ersten Vokals bei 263 Hz und die des zweiten Vokals bei 305 Hz. In beiden Parametern weist also die zweite Silbe in beiden Typen von Familiarisierungsstimuli einen höheren Wert auf als die erste. Demnach ist die zweite Silbe jeweils akustisch salienter als die erste, was dafür spricht, dass sowohl die jambischen Nomen als auch die Artikel-Nomen-Phrasen metrisch einem schwach-stark, also einem jambischen Muster folgen.

Auffällig ist jedoch, dass in beiden Parametern die Differenz zwischen den Werten bei den Nomen größer ist als bei den Phrasen. So sind die Vokale der zweiten Silbe der Nomen durchschnittlich 124 ms länger als die der ersten Silbe, bei den

Phrasen beträgt diese Differenz lediglich im Schnitt 80 ms. Die Vokale der zweiten Silbe der Nomen zeigen im Schnitt eine um 87 Hz höhere maximale Grundfrequenz als die der ersten Silbe, bei den Phrasen beträgt diese Differenz lediglich 42 Hz. Dieses Muster deutet daraufhin, dass die Prominenzunterschiede in den beiden Silben bei den Nomen vermutlich höher sind als bei den Phrasen. Nach Cutler (1993) sind für die Wahrnehmung von Betonung nicht die absoluten Werte der akustischen Korrelate innerhalb einer Silbe relevant, sondern Betonung ist eine relationale Eigenschaft, die zwischen den Silben einer Äußerung festzumachen ist. Daraus folgt, dass für die Wahrnehmung der Familiarisierungsstimuli als Einheiten bzw. für ihre Segmentierung wahrscheinlich die Differenzen zwischen den beiden Silben ausschlaggebend sind. Möglicherweise wiesen in unserem Material nur die Nomen so starke Differenzen auf, dass sie von den jüngeren Kindern als Segmentierungshinweis genutzt wurden, während die schwächeren Unterschiede der Phrasen für eine Segmentierung nicht ausreichten. Die Stichhaltigkeit dieser Erklärungsmöglichkeit muss jedoch in weiteren Experimenten überprüft werden.

Insgesamt bleibt jedoch festzuhalten, dass deutschsprachige Säuglinge im Alter ca. 11 bis 12 Monaten, auf Artikel-Nomen Sequenzen anders reagieren als auf jambische zweisilbige Nomen, deren erste Silbe kein eigenständiges Wort darstellt. Wie dargelegt, vermuten wir, dass dies darauf beruht, dass die Kinder den Artikel als eigenständige Einheit erkennen und diese Einheit aus dem Kontext herauslösen. In dieser Hinsicht liefert ihnen das Erkennen der Form Information über potentielle Wortgrenzen im Signal.

6.3.1.5.3 Experiment 8: Erkennung einer vertrauten Phonemsequenz oder eines Funktionswortes?

Die Frage, die nach den Ergebnissen des gerade beschriebenen Experiments noch offen bleibt, ist die nach der Art der mentalen Repräsentationen, über die die Kinder im Alter von 11 Monaten verfügen. Konkreter ausgedrückt heißt dies, erkennen die Kinder in den Artikelformen der Phrasen lediglich eine vertraute Phonemsequenz, die als Phonemmuster bereits mental gespeichert ist und über die im weiteren Verlauf des Spracherwerbs weitere Information gesammelt wird. Oder verfügen die Kinder bereits

über eine Art lexikalische Repräsentation über die Elemente, in der neben Wissen über den phonologischen Aufbau einer Einheit auch schon ansatzweise Wissen über die Funktion oder die Kontexte dieser Elemente enthalten ist. Dieser Frage wurde in einem weiteren Experiment nachgegangen, in dem die Artikelformen der Familiarisierungssequenz durch ein anderes hochfrequentes Funktionswort ersetzt wurden, das jedoch keine grammatisch wohlgeformte Phrase in Verbindung mit dem folgenden Nomen bildete. Wenn Kinder dieses Alters lediglich hochfrequente Phonemmuster erkennen und aus dem Input extrahieren, sollten sie sich in diesem Fall genauso verhalten wie im Fall der Phrasenfamiliarisierung und auf die starken Silben in den Textpassagen reagieren.

Probanden

Die Untersuchungsgruppe umfasste 26 Kinder mit einem durchschnittlichen Alter von 11 Monaten und 8 Tagen. Die Altersspanne reichte von 10 Monaten und 16 Tagen bis zu 12 Monaten und 19 Tagen. 24 weitere Kinder wurden getestet, ihre Daten konnte jedoch aufgrund von Experimentabbrüchen oder zu kurzen Orientierungszeiten nicht in die Auswertung einbezogen werden. Alle Kinder wuchsen in deutschsprachigen, monolingualen Elternhäusern auf und waren termgeboren.

Untersuchungsmaterial

Das Stimulusmaterial für die Familiarisierungsphase umfasste vier verschiedene Funktionswort-Nomen Sequenzen. Bei der Auswahl der Funktionswörter für das Experiment waren die folgenden Kriterien ausschlaggebend: Erstens sollten die Funktionswörter in ihrer Wortfrequenz eine möglichst gute Vergleichbarkeit mit den im vorherigen Experiment verwendeten Artikelformen haben. Zweitens durften sie mit dem folgenden Nomen keine grammatische Phrase bilden. Drittens sollten sie, wie die im vorherigen Experiment verwendeten Artikelformen, aus einer geschlossenen Silbe bestehen. Die Auswahl wurde anhand eines 15.000 Wörter umfassenden Korpus kindgerichteter Sprache getroffen. Nach diesem Korpus erfüllten die beiden Pronomen *was* und *wir* die Auswahlkriterien am besten.

Diese Funktionswörter wurden mit den gleichen einsilbigen Nomen kombiniert, die auch in den Experimenten 6 und 7 verwendet worden waren: *wir Kahn*, *wir Mohn*,

was Ziel, was Tor. Diese Funktionswort-Nomen Sequenzen wurden von der gleichen Sprecherin, die das Stimulusmaterial für die vorherigen Experimente eingesprochen hatte, aufgenommen. Sie erhielt die Anweisung die Funktionswort-Nomen Sequenzen in genau derselben Weise zu produzieren, wie eine Artikel-Nomen Sequenz, wiederum in einer kindgerichteten Sprache mit variabler Intonation.

Aus dieser Aufnahme wurden vier Familiarisierungssequenzen hergestellt, die bei einer Länge von ca. 30 s jeweils 26 Wiederholungen einer Funktionswort-Nomen Sequenz enthielten. In der Testphase wurden dieselben vier Textpassagen wie in den Experimenten 6 und 7 verwendet.

Durchführung

Die Durchführung des Experiments war die gleiche wie bei den vorherigen Experimenten. Wieder wurde den Kindern in der Familiarisierungsphase zwei der vier Sequenzen geboten, bis sie jede der beiden Funktionswort-Nomen Sequenzen für mindestens 30 Sekunden gehört hatten. Dabei wurden jeweils nur solche Sequenzen miteinander kombiniert, die zwei verschiedene Funktionswörter enthielten. Anschließend hörten die Kinder alle vier Textpassagen jeweils viermal.

Ergebnisse und Diskussion

Für die Analyse der Ergebnisse wurden zunächst für jedes Kind die Orientierungszeiten für die einzelnen Textpassagen ermittelt. Daraus wurden die Mittelwerte für die Textpassagen mit familiarisierter starker Silbe und für die Textpassagen ohne familiarisierte Silbe berechnet. Im Durchschnitt hörten die Kinder für 6,59 s den Textpassagen mit den familiarisierten starken Silben zu und für 7,07 s den Textpassagen ohne die familiarisierten starken Silben. Der t-Test für abhängige Stichproben zeigte, dass dieser Unterschied statistisch nicht signifikant war ($t(df = 25) = 1,41; p = 0,17$).

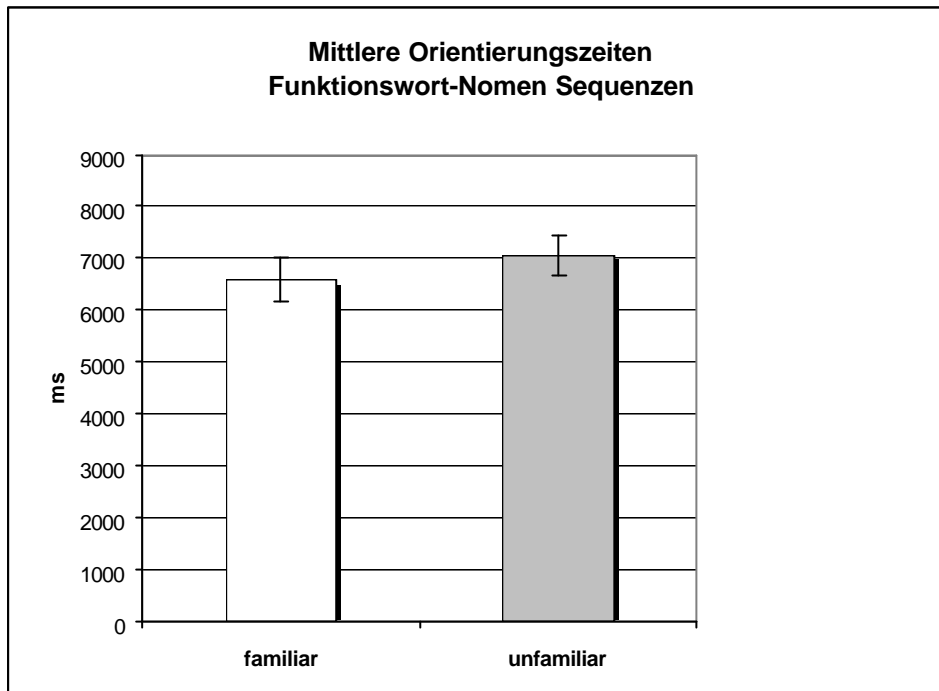


Abbildung 8: Mittlere Orientierungszeiten für Texte mit und Texte ohne familiarisierte starke Silbe aus Experiment 8

Dieses Ergebnis deutet darauf hin, dass die 10 bis 12 Monate alten Kinder, die mit den ungrammatischen Funktionswort-Nomen Sequenzen familiarisiert wurden, die starken Silben in den Textpassagen nicht wiederentdeckten. Damit verhalten sie sich eher wie ihre Altersgenossen aus Experiment 6, die mit den jambischen Nomen familiarisiert wurden, als wie die Kinder, die mit den Nominalphrasen familiarisiert wurden. Dies könnte darauf hindeuten, dass das Auftreten der Funktionswörter in einem ungrammatischen Kontext, ihre Erkennung als eigenständiges Element verhindert hat. Möglicherweise wird die schwache Silbe in diesem Fall zusammen mit der folgenden starken Silbe zu einer Einheit integriert. Die Berücksichtigung solcher kontextueller Merkmale ist durchaus wichtig, um den sprachlichen Input korrekt zu segmentieren, denn in vielen Fällen sind in einer gegebenen Wortform andere Wortformen enthalten, wie beispielsweise das Pronomen *er* im Nomen *Erbe*. Hier kann der Kontext zu einer korrekten Analyse des Inputs beitragen. Darüber hinaus unterstützt das Ergebnis von Experiment 8 den Befund der vorherigen Experimente, dass für Kinder im Alter von ca. 11 Monaten Artikelformen bereits einen besonderen Status haben: sie werden anders

analysiert als unbetonte Silben eines jambischen Wortes und wie in Experiment 8 gezeigt, auch anders als Funktionswörter mit vergleichbarer Wortfrequenz, wenn diese in einem ungrammatischen Kontext verwendet werden. Diese Ergebnisse lassen darauf schließen, dass Kinder gegen Ende des ersten Lebensjahres erstes Wissen über den Aufbau syntaktischer Phrasen in ihrer Zielsprache erworben haben.

6.3.1.6 Nutzung von Funktionswörtern für die syntaktische Klassifizierung

Bislang ist wenig darüber bekannt, auf welche Weise Kinder neue Wörter, die sie in ihr mentales Lexikon integrieren, syntaktisch klassifizieren. Wie in der Einleitung dargelegt, ist im Rahmen verschiedener Ansätze vorgeschlagen worden, dass Kinder semantische und distributionelle Merkmale zur syntaktischen Klassifizierung von Wörtern heranziehen. Dass Kinder auch formale Eigenschaften der Wörter zu deren Klassifizierung nutzen ist bislang so gut wie nicht untersucht. Eine Arbeit, die dieser Frage nachgeht, wurde von Cassidy und Kelly (1991) vorgelegt. Sie untersuchten, ob drei- bis vierjährige englische Kinder eine Assoziation zwischen der Länge eines Wortes, gemessen an der Silbenzahl, und seiner syntaktischen Kategorie herstellen. Dabei wurden den Kindern ein- und dreisilbige Pseudowörter in Zusammenhang mit einer Aktion präsentiert. Es wurde ihnen erklärt, dass diese Wörter in einer anderen Sprache entweder ein Objekt der Aktion oder die Aktion selbst bezeichneten. Die Kinder sollten schätzen, worauf das jeweilige Pseudowort referierte. Es zeigte sich dabei ein Zusammenhang zwischen der Länge des Wortes und der angenommenen Wortklasse: längere Wörter wurden häufiger als Objektbezeichnung klassifiziert als kürzere.

Hier soll der Frage weiter nachgegangen werden, inwieweit Kookkurrenzmuster zwischen Funktionswörtern und Inhaltswörtern den Kindern helfen können, die syntaktische Klasse der Inhaltswörter, die gemeinsam mit den Funktionswörtern auftreten, zu bestimmen. Dabei ist beispielsweise die Information relevant, dass einem Artikel mit einer relativ hohen Wahrscheinlichkeit ein Nomen folgt. Eine relativ frühe Fähigkeit, gerade die Funktionswörter im Input zu identifizieren und deren syntaktische Eigenschaften zu erwerben, könnte dann wiederum zum Lexikonerwerb beitragen, indem das Kind seine Kenntnisse über die distributionellen Merkmale nutzt, um die

Wortklasse neu gelernter Wörter zu ermitteln. Die Berücksichtigung distributioneller Information sollte die zuverlässigste Methode sein, die Klasse eines noch unbekanntes Wortes zu erkennen, da Wortklassen anhand ihrer distributionellen Eigenschaften definiert sind. Es wurde von vielen Autoren vorgeschlagen, dass Kinder diese Information zur Klassifizierung von Wörtern und eventuell auch größeren Einheiten wie syntaktischen Phrasen nutzen (z.B. Gerken, 2001; Valian & Coulson, 1988), empirische Evidenz für diese Hypothese fehlt jedoch bislang fast vollständig.

Die einzigen Hinweise darauf, dass Kinder die Information von Funktionswörtern zur Klassifikation anderer Wörter heranziehen, stammt aus einer Arbeit von Katz, Baker und Macnamara (1974). Sie stellten fest, dass 17 Monate alte Kinder in Abhängigkeit davon, ob ihnen ein neues Wort als Bezeichnung für eine Puppe zusammen mit einem Artikel oder ohne Artikel präsentiert wurde, dieses Wort entweder als Eigenname oder als Gattungsname interpretierten.

In einem eigenen Experiment sollte der Frage der Ausnutzung von Kookkurrenzbeziehungen zwischen Funktionswörtern und Inhaltswörtern detaillierter nachgegangen werden. Dazu wurden die Kinder mit Pseudowörtern familiarisiert, die entweder gemeinsam mit einem Artikel oder mit einem Personalpronomen auftraten. Ansonsten wiesen die Pseudowörter keinerlei Merkmal auf, die sie als Mitglieder einer bestimmten Wortklasse auswiesen. In der Testphase wurden den Kinder Textpassagen präsentiert, in denen die Pseudowörter entweder als Nomen oder als Verben verwendet wurden. Die Frage war, ob die Kinder in Abhängigkeit davon, ob sie das Pseudowort zuvor zusammen mit einem Artikel oder zusammen mit einem Personalpronomen gehört hatten, unterschiedlich auf die Nomen- und die Verbverwendung der Pseudowörter in den Textpassagen reagieren würden. Dies würde dafür sprechen, dass sie die Funktionswörter, mit denen die Pseudowörter während der Familiarisierung kombiniert waren, zur syntaktischen Klassifikation der Pseudowörter heranziehen. Entsprechend der Befunde von Katz et al. (1974), dass 17 Monate alte Kinder das Vorhandensein bzw. Nicht-Vorhandensein eines Artikels für die Subkategorisierung von Nomen verwenden, sollten etwas jüngere Kinder als in dieser Studie untersucht werden. Deshalb wurden Kinder im Alter von ca. 15 Monaten in die Studie einbezogen.

6.3.1.6.1 Experiment 9: Nutzung von Funktionswörtern zur syntaktischen Kategorisierung bei 15 Monate alten Kindern

Probanden

Insgesamt wurden 69 Kinder im Alter zwischen 14 und 16 Monaten in diese Untersuchung einbezogen. Alle Kinder wuchsen in monolingualen, deutschsprachigen Elternhäusern auf und waren termgeboren. Die Daten von 21 dieser Kinder konnten nicht in die Datenanalyse einbezogen werden: bei 18 Kindern musste das Experiment abgebrochen werden, 1 Kind wies zu kurze Orientierungszeiten von unter 3 Sekunden für eine der Experimentalbedingungen auf und bei zwei Kindern konnte das Experiment wegen technischer Probleme nicht zu Ende geführt werden. Von den übrigen Kindern gehörten 24 zur Gruppe, die mit den Artikel-Pseudowort-Sequenzen familiarisiert worden waren. Diese Kinder hatten ein durchschnittliches Alter von 15 Monaten und 18 Tagen. Die Altersspanne lag zwischen 14 Monaten und 4 Tagen und 16 Monaten und 24 Tagen. In dieser Gruppe waren 11 Mädchen und 13 Jungen. In der Gruppe, die mit den Pronomen-Pseudowort-Sequenzen familiarisiert wurden, befanden sich ebenfalls 24 Kinder. Diese wiesen ein durchschnittliches Alter von 15 Monaten und 21 Tagen auf, mit einer Spanne von 14 Monaten und 7 Tagen bis 16 Monaten und 26 Tagen. In dieser Gruppe befanden sich ebenfalls 11 Mädchen und 13 Jungen.

Untersuchungsmaterial

Zunächst wurden zwei verschiedene einsilbige Pseudowörter konstruiert (glamm, pronk). Das Hauptaugenmerk wurde dabei darauf gelegt, dass diese Pseudowörter keine Merkmale trugen, die eine Klassifikation als Verb oder als Nomen nahegelegt hätten. Dies beeinflusste unter anderem die Wahl der Vokale in den Pseudowörtern. Verben weisen im Deutschen im allgemeinen eine Flexionsendung auf. Die verwendeten Formen sollten jedoch kein flexionsähnliches Element haben, da dieses eventuell zur Kategorisierung herangezogen werden könnte. Eine systematische Ausnahme zu dieser generellen Flektiertheit von Verben bilden die sogenannten starken Verben. Diese haben im Präterium in der ersten und dritten Person Singular keine Flexionsendung. Da in diesen Fällen die Ablautung als formale Markierung für das Präterium eingesetzt wird, weisen diese Formen jedoch typische Vokale auf, die auch in den Pseudowörtern verwendet wurden.

Für die Familiarisierungsphase wurden die Pseudowörter zum einen mit dem indefiniten Artikel zu den Sequenzen *ein glamm* und *ein pronk* kombiniert. Zum anderen wurden sie mit dem Personalpronomen *sie* zu den Sequenzen *sie glamm* und *sie pronk* zusammengefügt. Für jedes Pseudowort wurden zwei aus sechs einfachen Sätzen bestehende Textpassagen konstruiert. In jedem Satz der Textpassage kam das jeweilige Pseudowort einmal vor. In einer der beiden Passagen wurde das Pseudowort ausschließlich in Kontexten verwendet, in denen eindeutig ein Nomen erforderlich war, in der anderen Textpassage wurde das Pseudowort in eindeutigen Verbkontexten verwendet. In keiner der Textpassagen wurde eins der Funktionswörter, die in den Familiarisierungssequenzen vorkamen benutzt.

Das gesamte Untersuchungsmaterial wurde von einer Muttersprachlerin des Deutschen gesprochen. Sie hatte die Anweisung, die Stimuluswörter und -texte in einer kindgerichteten Weise mit variabler Intonation zu produzieren. Anschließend wurde das Material digitalisiert. Für jede der vier Textpassagen wurde eine einzelne Datei angelegt. Die Passagen wiesen eine durchschnittliche Dauer von 18 s auf (*glamm*-Nomen: 18,63 s; *pronk*-Nomen: 17,21 s; *glamm*-Verb: 17,59 s; *pronk*-Verb: 18,61 s). Für jede der Familiarisierungssequenzen wurde je eine Datei hergestellt, die jeweils 26 Token einer Artikel-Pseudowort-Sequenz bzw. 27 Token einer Pronomen-Pseudowort-Sequenz enthielt. Zwischen den einzelnen Token war eine Pause von 600 ms. Insgesamt wiesen die Familiarisierungssequenzen eine Länge von ca. 30 Sekunden auf.

Untersuchungsdurchführung

Es wurde das *Headturn Preference* Paradigma mit Familiarisierungsphase wie beschrieben verwendet. In der Familiarisierungsphase hörten die Kinder zwei der insgesamt vier Familiarisierungssequenzen und zwar jeweils die zwei mit dem gleichen Funktionswort, also beide Artikel-Sequenzen oder beide Pronomen-Sequenzen. In der Testphase wurden allen Kindern alle vier Textpassagen insgesamt viermal präsentiert.

Ergebnisse und Diskussion

Für jedes Kind wurden die Orientierungszeiten für jede der vier Textpassagen ermittelt. Anschließend wurden die mittleren Orientierungszeiten für die Textpassagen, in denen

das Pseudowort als Nomen verwendet wurde und für die Textpassagen, in denen das Pseudowort als Verb verwendet wurde berechnet.

Die Kinder, die mit den Artikel-Pseudowort-Sequenzen familiarisiert worden waren zeigten durchschnittliche Orientierungszeiten von 6,95 s für die Nomen-Passagen und von 7,90 s für die Verbpassagen. Dieser Unterschied erwies sich als statistisch signifikant ($t(df = 23) = 2,86; p < 0,01$). Die durchschnittlichen Orientierungszeiten der Kinder, die mit den Pronomen-Pseudowort-Sequenzen familiarisiert worden waren, betragen 6,22 s für die Nomen-Passagen und 6,14 s für die Verb-Passagen. Der Unterschied war statistisch nicht bedeutsam ($t(df = 23) = 0,21; p = 0,83$).

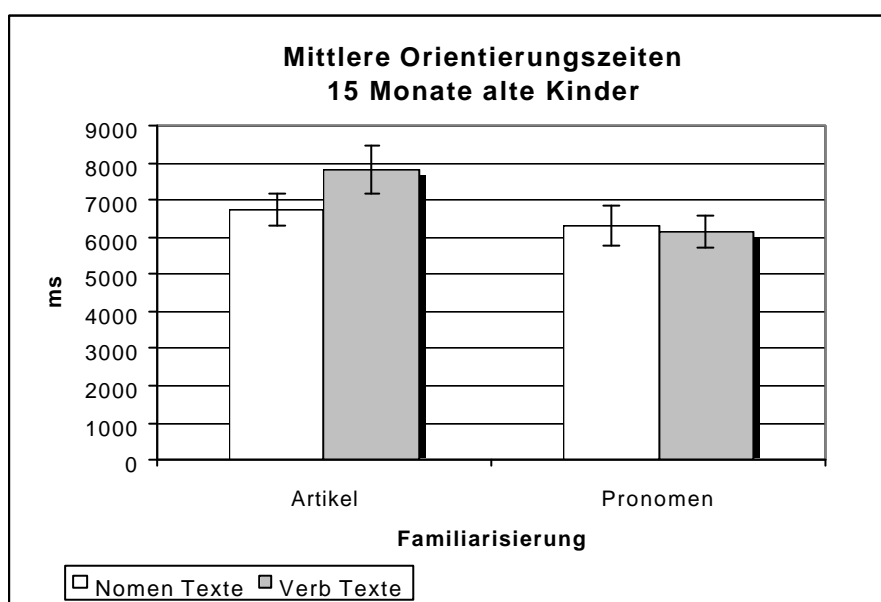


Abbildung 9: Mittlere Orientierungszeiten für Nomen- und Verb-Passagen aus Experiment 9

Dieses Ergebnis deutet darauf hin, dass Kinder im Alter von ungefähr 15 Monaten über Wissen über Kookkurrenzmuster verschiedener Wortklassen verfügen und dass sie dieses Wissen einsetzen können, um unbekannte Wörter syntaktisch zu kategorisieren. Zumindest die Befunde bei den Kindern, die mit den Artikel-Pseudowort-Sequenzen familiarisiert wurden, weisen klar in diese Richtung. Offensichtlich erkannten sie die Form des indefiniten Artikels in den Familiarisierungssequenzen und zogen ihr vorhandenes Wissen über die üblichen syntaktischen Kontexte, in denen die

Artikelform auftritt, heran, um zu schließen, dass die dem Artikel folgende Phonemkette zu einer Wortklasse gehört, die wir als Nomen bezeichnen. Nur so ist zu erklären, dass diese Kinder in der Testphase systematisch unterschiedlich auf die Nomen- und die Verb-Passagen reagierten. Hier zeigt sich jedoch über die Klassifikation des Pseudowortes hinaus eine weitere Fähigkeit: die Kinder müssen in der Lage gewesen sein zu erkennen, dass sich der Gebrauch der Pseudowörter in den Nomen- und den Verb-Passagen unterschied. Da in den Textpassagen weder der indefinite Artikel noch die in der Familiarisierung verwendete Pronomenform enthalten war, bedeutet dies, dass die Kinder bereits in der Lage sein müssen, bestimmte Typen syntaktischer Kontexte unabhängig davon, welche Art von lexikalischen Elementen den Kontext bilden, als kategoriell gleich zu identifizieren, d. h. sie sind in der Lage, auch Phrasentypen zu erkennen.

Die Annahme, dass die unterschiedlichen Reaktionen der Artikel-Familiarisierungs-Gruppe auf die Textpassagen tatsächlich etwas damit zu tun hatte, dass die Kinder während der Familiarisierung durch die Kontexte Wortklasseninformation über das Pseudowort erhielten, wird auch durch die Ergebnisse der Kinder, die mit den Pronomen-Pseudowort-Sequenzen familiarisiert wurden, gestützt. Wie gesagt, fanden sich bei diesen Kindern keine systematischen Unterschiede in den Orientierungszeiten für die Nomen- und die Verb-Passagen. Dies zeigt, dass das Reaktionsmuster der Artikel-Familiarisierungs-Gruppe nicht auf einer generellen Präferenz für die Verb-Passagen beruht. Dies wäre durchaus möglich gewesen, da die Nomen- und Verb-Passagen nicht nur strukturell im Gebrauch der Pseudowörter sondern auch inhaltlich sehr unterschiedlich waren.

Auf der anderen Seite zeigen die Befunde für die Pronomen-Familiarisierungs-Gruppe aber auch, dass die Aussage, Kinder dieses Alters nutzten Kookkurrenzen zwischen Funktionswörtern und Inhaltswörtern für die syntaktische Klassifikation zunächst auf die Artikel-Nomen-Beziehung eingeschränkt werden muss. Es gibt kein Indiz, dass die Pronomen in ähnlicher Weise wie die Artikel für die Klassifizierung der Pseudowörter ausgenutzt wurden. Ein Faktor, der möglicherweise zu diesem Unterschied zwischen Pronomen und Artikeln beiträgt, könnte darin bestehen, dass der Artikel ein weitaus reliablerer Prädiktor für die Wortklasse des folgenden Wortes

darstellt, als das Pronomen, d.h. möglicherweise folgt im Input des Kindes tatsächlich einem Artikel weitaus häufiger ein Nomen als einem Pronomen ein Verb.

Um diese Annahme zu überprüfen wurden Artikel- und Pronomenkontexte in einem Korpus deutscher kindgerichteter Sprache analysiert. Das Korpus umfasst 16 Aufnahmen von Spontansprache, die eine Mutter während natürlicher alltäglicher Situationen an ihr Kind gerichtete hatte. Die Aufnahmen waren gemacht worden, als das Kind zwischen 18 und 27 Monaten alt war. Das Korpus umfasste ca. 15.000 Wörter. Aus diesem Korpus wurden ausschließlich die Artikel- und Pronomenform, die im Experiment verwendet worden waren (d. h. *ein* und *sie*), extrahiert und die Wortklasse des direkt folgenden Wortes bestimmt. Für den indefiniten Artikel fanden sich insgesamt 179 Verwendungen innerhalb des Korpus. In 71% dieser Fälle war das folgende Wort tatsächlich ein Nomen. Das Personalpronomen wurde insgesamt 64 mal verwendet. Dabei folgte nur in 31% der Fälle ein Verb.

Diese Zahlen deuten darauf hin, dass die Kontexte, in denen Artikel auftreten, sehr viel restringierter sind als die Kontexte, in denen ein Pronomen auftritt. Dies liegt sicherlich zum einen darin, dass die Pronomenform *sie* kasusambig, d. h. sie ist sowohl Nominativform als auch Akkusativform und kann damit sowohl als Subjekt als auch als Objekt verwendet werden. Ein wichtiger Punkt für die größere kontextuelle Vielfalt des Pronomens dürfte zudem auch die relativ freie Wortstellung des Deutschen sein. Das Subjekt muss eben nicht immer dem Verb vorangehen, so dass selbst in Fällen, in denen das Pronomen als Subjekt verwendet wird, es nicht notwendigerweise vor dem Verb erscheinen muss.

Die Daten der 15 Monate alten Kinder reflektieren diese Unterschiede in den kontextuellen Restriktionen. Vor dem Hintergrund der berichteten äußersten Sensitivität von Kindern für die Häufigkeit von Kookkurrenzmustern in ihrem sprachlichen Input ist die Annahme nicht unplausibel, dass die Kinder im Alter von 15 Monaten die regelmäßige Beziehung zwischen Artikel und Nomen erkannt haben, jedoch keine ähnlich deutliche Beziehung zwischen Pronomen und Verben herstellen konnten.

Es ist jedoch auch noch eine andere Erklärung dafür denkbar, dass das Pseudowort in Verbindung mit dem Pronomen nicht als Verbform klassifiziert wurde. Wie eingangs schon erwähnt, verfügt die Mehrzahl der deutschen Verbformen über eine Flexionsendung. Die verwendeten flexionslosen Formen, sind möglicherweise keine

typischen Verbformen, obwohl die flexionslose Form in den starken Verben ja durchaus existent ist. Allerdings ist diese Klasse klein und nicht mehr produktiv. Auf der anderen Seite gehören jedoch gerade eine Vielzahl von hochfrequenten Verben wie *gehen*, *fahren*, *essen* etc. in diese Klasse. Nach unseren bisherigen Überlegungen wäre es möglich, dass deutschlernende Kinder eine starke Assoziation zwischen der Klasse der Verben und dem Vorhandensein einer Flexionsendung gebildet haben, so dass die verwendeten flexionslosen Formen aufgrund dieser Tatsache nicht als Verb angesehen werden. Eine Entscheidung zwischen diesen beiden Erklärungen ist anhand vorhandener Daten bislang nicht möglich.

Festzuhalten bleibt als wichtiges Ergebnis, dass Kinder im Alter von 15 Monaten prinzipiell in der Lage, Kookkurrenzen zwischen verschiedenen Wortklassen für die syntaktische Klassifizierung zu nutzen. Darüber hinaus verfügen sie bereits über spezifischere lexikalische Einträge über Artikel, in denen neben Information über die Wortform auch bereits syntaktische Eigenschaften gespeichert sein müssen. Es stellt sich nun die Frage, ob diese Fähigkeit auch bereits bei jüngeren Kindern vorhanden sind. In Experiment 6 zeigte sich ja, dass Kinder im Alter von ca. 11 Monaten Artikelformen bereits als vertraute Phonemsequenzen erkennen. Es stellt sich also noch einmal die Frage, ob auch jüngere Kinder schon über speziellere lexikalische Repräsentationen für diese Formen verfügen. Aus diesem Grund wurde Experiment 9 mit einer Gruppe jüngerer Kinder wiederholt.

6.3.1.6.2 Experiment 10: Nutzung von Funktionswörtern für die syntaktische Klassifizierung bei Kindern im Alter von 12 und 13 Monaten

Probanden

In dieser Altersgruppe wurden insgesamt 70 Kinder getestet. Bei 18 Kindern musste das Experiment abgebrochen werden, drei Kinder wurden wegen zu kurzer Orientierungszeiten aus der Analyse ausgeschlossen und eine Durchführung konnte wegen technischer Probleme nicht beendet werden. 24 der Kinder wurden der Nomen-Familiarisierungs-Gruppe zugeordnet. Diese Gruppe wies ein durchschnittliches Alter von 12 Monaten und 21 Tagen auf bei einer Spannbreite von 12 Monate 2 Tage bis 13 Monate 23 Tage. Die Gruppe umfasste 12 Mädchen und 12 Jungen. Der Verb-

Familiarisierung wurden ebenfalls 24 Kinder zugewiesen. Das mittlere Alter in dieser Gruppe lag bei 12 Monaten 28 Tagen mit einer Spannbreite von 12 Monaten 2 Tagen bis 13 Monate und 30 Tage. In dieser Gruppe befanden sich 8 Mädchen und 16 Jungen.

Untersuchungsmaterial und Untersuchungsdurchführung

Material und Durchführung waren identisch zu Experiment 10.

Ergebnisse und Diskussion

Wieder wurden für jedes Kind die Mittelwerte für die vier Textpassagen berechnet und der Mittelwert für die Verb- und die Nomenpassagen gebildet. Die Kinder der Nomen-Familiarisierung wiesen durchschnittliche Orientierungszeiten von 6,41 s für die Nomen-Passagen und von 6,54 s für die Verbpassagen auf. Der Unterschied ist

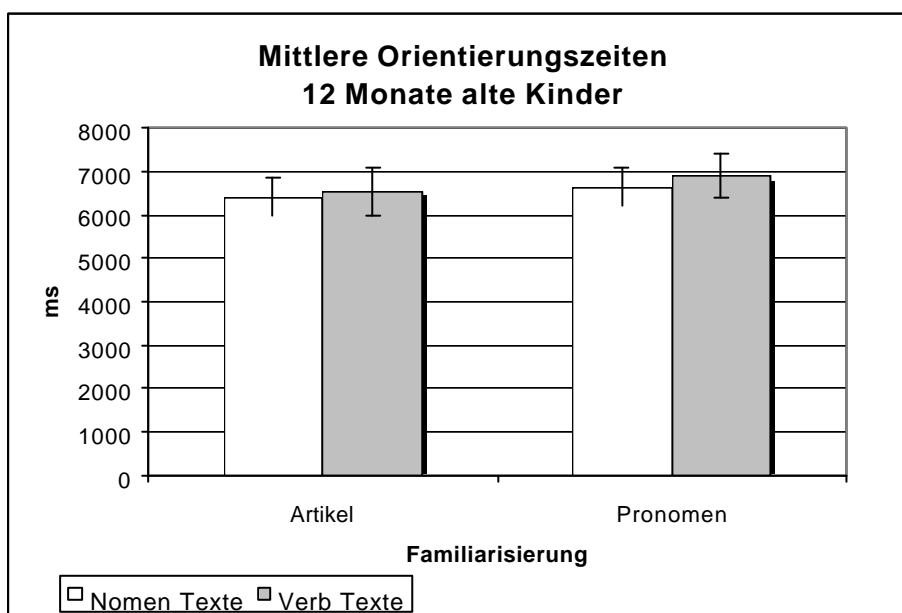


Abbildung 10: Mittlere Orientierungszeiten für Nomen- und Verb-Passagen aus Experiment 10

statistisch nicht signifikant ($t(df = 23) = 0,35; p = 0,73$). Bei den Kindern der Verb-Familiarisierung zeigten sich mittlere Orientierungszeiten von 6,66 s für die Nomen-

Passagen und von 6,88 s für die Verb-Passagen. Auch dieser Unterschied ist statistisch nicht bedeutsam ($t(df = 23) = 0,44; p = 0,66$).

Im Gegensatz zu den älteren Kindern findet sich in dieser Altersgruppe also noch kein Unterschied in den Ergebnissen der beiden Familiarisierungsgruppen. Weder die Kinder der Nomen-Familiarisierung noch die Kinder der Verb-Familiarisierung zeigten in der Testphase Unterschiede in den Reaktionen auf die verschiedenen Textpassagen. Dieses Ergebnismuster kann darauf beruhen, dass die Kinder im Alter von 12 bis 13 Monaten die durch den infiniten Artikel gegebene Information nicht zur Klassifizierung des folgenden Pseudowortes nutzen. Es könnte jedoch auch dadurch zustande gekommen sein, dass die Kinder diesen Alters noch nicht in der Lage sind, die dem Pseudowort zugeordnete Wortklasse über verschiedene Kontexte zu generalisieren, d. h. zu erkennen, dass in den Textpassagen zum einen Kontexte für ein Nomen auftraten und zum anderen Kontexte für ein Verb. Zwischen diesen beiden Erklärungsmöglichkeiten lässt sich anhand der vorliegenden Daten nicht entscheiden.

6.4. Der Beitrag segmental-sequentieller Information zur Lösung des Segmentierungs- und Klassifizierungsproblems

Insgesamt zeigen die in diesem Kapitel referierten Befunde ein differenziertes Bild der Fähigkeiten von Säuglingen ihren sprachlichen Input auf segmental-distributioneller Ebene zu analysieren, sich Wissen über Merkmale der Zielsprache anzueignen und dieses Wissen für die weitere Analyse von Äußerungen und damit für den Ausbau des eigenen Wissens heranzuziehen. Im einzelnen haben wir gesehen, dass quasi von Geburt an die Fähigkeit vorhanden ist, zwischen segmental verschiedenen sprachlichen Reizen zu diskriminieren. Dabei kann bereits von konkreten akustischen Merkmalen des Reizes, die durch Sprechercharakteristika oder der Variation der phonetischen Realisierung innerhalb eines Sprechers bedingt sind, abstrahiert werden. Die gezeigten Diskriminationsfähigkeiten belegen auch, dass schon von Geburt an - zumindest kurzfristig - sprachliche Reize mental repräsentiert werden können, die bis zu einem gewissen Grad auch schon segmental spezifiziert sind. Dieser Schluss ergibt sich daraus, dass in den gängigen Paradigmen, die zur Untersuchung der Lautdiskrimination

verwendet werden, Gedächtnisleistungen involviert sind, da die zu diskriminierenden Reize immer nacheinander und nicht parallel präsentiert werden können. Zudem können innerhalb des ersten Lebensjahres offensichtlich auch bereits längerfristige Repräsentationen für solche sprachlichen Sequenzen gebildet werden, die der Säugling in seinem Input sehr häufig hört, wie es für den eigenen Namen gezeigt wurde.

In der zweiten Hälfte des ersten Lebensjahres lassen sich weitreichende Fähigkeiten zur distributionellen Analyse von sprachlichen Äußerungen nachweisen, die vor allen Dingen in Bezug auf die Kookkurrenzen von Silben gezeigt wurden, und die von den Kindern auch bereits genutzt werden, um Einheiten im sprachlichen Input zu finden. Es ist bislang nichts darüber bekannt, ob sich diese Analysemechanismen erst während des ersten Lebensjahres entwickeln oder ob auch diese bereits von Geburt an vorhanden sind. In der zweiten Hälfte des ersten Lebensjahres sehen wir zudem bereits weitreichende Kenntnisse über distributionelle Merkmale der Zielsprache: die Kinder haben Kenntnis über das segmentale Inventar, über phonotaktische Beschränkungen und über die Häufigkeit von Konsonantenabfolgen in bestimmten Wortpositionen. Mit Ende des ersten Lebensjahres verfügen sie offensichtlich auch schon über Langzeitrepräsentationen zumindest der phonologischen Form von Artikeln und etwas später, im Alter von ca. 15 Monaten auch bereits über syntaktisches Wissen über Artikel. All dieses Wissen wird offensichtlich sobald es verfügbar ist, vom Säugling für Segmentierungs- und Klassifikationsprozesse für seinen weiteren Input herangezogen.

Wie schon mehrfach erwähnt, ist der Aufbau und die Nutzbarkeit dieses segmental-sequentiellen Wissens jedoch eng an bereits vorhandenes Wissen geknüpft, so dass hier noch einmal die Frage diskutiert werden muss, inwieweit dieser Informationstyp tatsächlich im Sinne eines Bootstrapping-Mechanismus den Einstieg in den Erwerb auf höheren sprachlichen Ebenen ermöglichen kann. Eine reine *bottom-up* Nutzung ist lediglich für rein akustische Hinweise möglich. Für die von den Säuglingen herangezogenen Merkmale, ist eine reine *bottom-up* Funktionsweise nur in zwei Domänen denkbar: zum einen im Bereich der akustischen Hinweise, wo sich gezeigt hat, dass Säuglinge Koartikulationsmerkmale zur Strukturierung des Inputs heranziehen und diesen sogar stärker vertrauen als distributionellen Merkmalen. Die Kookkurrenzmuster sind der zweite Informationstyp, deren Beachtung ohne zusätzliches bereits vorhandenes Wissen über potentiellen Eigenschaften der zu

bestimmenden Einheiten, zu einer Analyse des sprachlichen Inputs beitragen kann. Es scheint fraglich, ob der sprachliche Input Kookkurrenzmuster in einer Dichte enthält, die dem menschlichen Verarbeitungssystem zugänglich ist. Aber die Befunde, nach denen Säuglinge in der Sprache häufige Lautkombinationen von selteneren und häufig wortintern auftretende Konsonantenfolgen von selten wortintern auftretenden Folgen unterscheiden können, deutet darauf hin, dass auch natürlichsprachlicher Input ausreichend Information für eine erfolgreiche Analyse zumindest bestimmter distributioneller Merkmale enthält. Wie erwähnt erleichtert gerade diese Sensivität für Kookkurrenzen auf der Phonemebene eine frühe Extraktion von häufigen, einsilbigen Wörtern aus dem Input.

Alle übrigen Informationstypen können nur auf der Basis bereits vorhandenen Wissens über die relevanten Einheiten genutzt werden. Allophonische Varianten können nur dann zur Segmentierung genutzt werden, wenn Wissen darüber vorliegt, dass wortinitial, wortfinal oder wortmedial ausschließlich bestimmte Allophone vorkommen. Wie schon erwähnt setzt die Erkennung phonotaktischer Regularitäten zumindest das Wissen über Silben als relevante Größe voraus. Die Berechnung der Häufigkeit wortinterner vs. wortübergreifender Konsonantenfolgen kann nur nach einer bereits erfolgten Wortsegmentierung vonstatten gehen. Allerdings zeigt sich, dass bereits Säuglinge offensichtlich alle Merkmale, die sich auch bei Erwachsenen als hilfreich für die Erkennung von Einheiten und Regularitäten im Input erwiesen haben, nutzen können.

Am Anfang dieses Kapitels wurde die Frage gestellt, ob das Einbeziehen segmental-distributionellen und lexikalischen Wissens zur Lösung von Problemen beitragen kann, die anhand prosodischer Information allein nicht gelöst werden können. Sicherlich können alle die in diesem Kapitel aufgezeigten Merkmale dazu beitragen, anfängliche - rhythmisch bedingte - Fehlsegmentierungen bei jambischen Wörtern zurückzudrängen und damit auch jambische Wörter als Einheit zu erkennen. Wie berichtet, bestehen offensichtlich gegen Ende des ersten Lebensjahres bei der Erkennung von jambischen Wörtern keine Probleme mehr. Da gerade in der zweiten Hälfte des ersten Lebensjahres eine zunehmende Fähigkeit zu beobachten ist, segmental-sequentielle Information in die Segmentierung einzubeziehen, ist die

Annahme, dass dies zur erfolgreichen Erkennung auch jambischer Wörter beiträgt, naheliegend.

Für die rhythmische Segmentierungsstrategie stellen neben jambischen Wörtern auch unbetonte Funktionswörter ein besonderes Problem dar. Diese sollten nach einer reinen rhythmusbasierten Strategie gar nicht als eigenständige Einheit erkannt werden können, sondern stets einem vorangehenden Wort als unbetonte Silbe angefügt werden. Auch hier kann das zunehmende Wissen über Kookkurrenzen und das Heranziehen dieses Wissens für die Segmentierung eine Lücke des prosodischen Bootstrapping füllen. Wie erwähnt sind gerade unbetonte Funktionswörter in der Sprache im allgemeinen hochfrequent, so dass sie als hochfrequentes Phonemmuster von einem distributionellen Analysemechanismus relativ schnell als unabhängige Einheit identifiziert werden sollten. Wie gesehen verfügen Kinder ja auch bereits gegen Ende des Lebensjahres über zumindest eine Repräsentation der lautlichen Form einiger dieser Elemente.

Bleibt die syntaktische Ebene. Die distributionelle Analyse liefert die zuverlässigere Information über die Zugehörigkeit eines Wortes zu einer syntaktischen Kategorie als prosodische Merkmale. Wie gesehen, sind die Korrelationen zwischen prosodischen Merkmalen und der Wortklasse zumindest in Bezug auf die Differenzierung verschiedener Klassen innerhalb der Inhaltswörter eher unscharf. Allerdings setzt die Nutzung von Kookkurrenzmustern für diesen Zweck offensichtlich schon etwas längere Lernprozesse voraus. In unseren Untersuchungen konnten wir sie für die Artikel-Nomen-Beziehung erst im Alter von 15 Monaten nachweisen und auch zunächst ausschließlich für diese Beziehung. Diese Überlegungen können auch auf größere syntaktische Einheiten übertragen werden. Es gibt meines Wissens keinerlei prosodische Hinweise, die eine Differenzierung zwischen verschiedenen Phrasentypen erlauben würden. Dies lässt sich jedoch ebenso wie für die Wortkategorisierung anhand der Beachtung der Information von Funktionswörtern erreichen: so klassifiziert beispielsweise der Artikel die gesamte syntaktische Phrase als Determinerphrase (vgl. Gerken, 1996).

Auch für die Erkennung von syntaktischen Grenzen innerhalb einer Äußerung können Funktionswörter ausgenutzt werden. Wie erwähnt bestehen vor allen Dingen deshalb keine isomorphen Beziehungen zwischen prosodischer und syntaktischer

Struktur, da nicht jede syntaktische Grenze auch prosodisch markiert sein muss. Da Funktionswörter häufig gerade an den Grenzen syntaktischer Phrasen auftauchen, könnten sie auch für die Identifizierung einer Phrasengrenze genutzt werden. Auf der anderen Seite wurde argumentiert, dass gerade diese exponierte Position an Phrasengrenzen in den Fällen, in denen diese prosodisch gut markiert sind, zu einer leichten Extraktion der unbetonten Funktionswörter aus dem Signal beitragen kann. Dieses letzte Beispiel macht deutlich, dass die verschiedenen prosodischen, segmentalen und lexikalischen Hinweise auf die Struktur einer Äußerung sich gegenseitig bedingen und offensichtlich in ihrer Interaktion die wertvollsten und reliabelsten Hinweise für die Analyse des sprachlichen Inputs liefern.

Unsere Befunde einer frühen Verarbeitung von Funktionswörtern und der Ausnutzung der darin enthaltenen Information in weiteren Analyseprozessen sprechen gegen die Annahme einer anfänglichen Priorität semantischer bzw. konzeptueller Kategorien und Relationen, wie sie in den in Kapitel 1 näher beschriebenen Ansätzen vorgeschlagen wurde (Schlesinger, 1982, 1988). Offensichtlich sind die Kinder bereits früh in der Lage auch sprachliche Kategorien zu bilden, die keinen deutlichen semantischen Kern aufweisen (vgl. Weissenborn, 2000). Nach unseren Befunden sind bereits 15 Monate alte Kinder, die sich im allgemeinen noch in der Einwortphase befinden, bereits in der Lage, neue lexikalische Einheiten allein nach rein formalen distributionellen Kriterien einer syntaktischen Kategorie zuzuordnen. Dieses Ergebnis spricht gegen die Annahme, dass die in diesem Alter oder kurze Zeit später auftauchenden ersten Mehrwortäußerungen ausschließlich semantische Relationen ausdrücken und ihnen nicht bereits eine syntaktische Struktur zugrundeliegt. Über unsere Untersuchungen hinaus findet sich auch in anderen Bereichen des Spracherwerbs Evidenz für die frühe Fähigkeit, Kategorien allein anhand formaler Kriterien zu bilden und formale Aspekte der Sprache wie beispielsweise Flexionsparadigmen zu erwerben (siehe die Arbeiten in Slobin, 1985). Dies gilt beispielsweise auch für das im allgemeinen semantisch hochgradig arbiträre Genussystem. Nach Levys (1983) Befunden fällt es Kindern leichter, ein System zu erwerben, in dem das Genus eines Nomens klar an seiner phonologischen Form ablesbar ist, als wenn semantische Kriterien des natürlichen Geschlechts eines Referenten in die Genuszuweisung einbezogen werden müssen. Ein ähnliches Muster zeigt Gordon (1985) für den Erwerb

der Differenzierung zwischen Massennomina und zählbaren Nomen. So sind bereits Zweijährige in der Lage, den Gebrauch des indefiniten Artikels zusammen mit einem Nomen als Indikator dafür zu interpretieren, dass es sich bei dem Nomen nicht um ein Massennomen handelt. Semantische Kriterien scheinen dagegen weniger beachtet zu werden.

Bei der nachgewiesenen Sensitivität für Funktionswörtern und dem bereits vorhandenen Wissen über diese Elemente und die Strukturen, in denen sie auftreten, stellt sich natürlich die Frage, warum gerade diese Elemente in den ersten Mehrwortäußerungen der Kinder fehlen. Zur Erklärung dieser Diskrepanz zwischen Produktionsmustern und den rezeptiven syntaktischen Fähigkeiten wurde von LouAnn Gerken ein metrischer Ansatz vorgeschlagen. Ausgehend von den Befunden, nach denen sich ein deutlicher Unterschied zwischen der Verwendung von funktionalen Elementen in der Sprachproduktion und deren Erkennung und Ausnutzung in der Rezeption zeigt (Gerken, Landau & Remez, 1990; Gerken & McIntosh, 1993; Shipley, Smith & Gleitman, 1969), vertritt Gerken die Auffassung, dass ausschließlich produktionsinterne und zwar nicht-syntaktische Faktoren für das Fehlen von Funktionswörtern in den frühen kindersprachlichen Äußerungen ursächlich sind. Sie vermutet, dass die frühe Sprachproduktion zunächst durch die Präferenz für ein bestimmtes metrisches Muster, nämlich das trochäische stark-schwach Muster geprägt ist. Die Bevorzugung dieses metrischen Musters führt dazu, dass Kinder solche unbetonten Silben auslassen, die nicht in das metrische Muster passen, also keiner betonten Silbe folgen, an die sie klitisiert werden können. Empirische Evidenz für diese Hypothese fand Gerken in verschiedenen Nachsprechaufgaben, in denen sich genau dieses Muster zeigte: unbetonte Elemente wurden von den Kindern mit höherer Wahrscheinlichkeit produziert, wenn sie einer betonten Silbe folgten (Gerken, 1991, 1994a, 1994b). Attraktiv an diesem Erklärungsansatz ist die Tatsache, dass er zwei zunächst nicht unbedingt assoziiert erscheinende Phänomene der frühen Kindersprache auf eine gemeinsame Ursache zurückführen kann: sowohl das Fehlen unbetonter Silben am Wortanfang als auch das Fehlen von Funktionselementen können auf diese metrische Präferenz zurückgeführt werden. Allerdings kann dieser Ansatz der beobachteten systematischen Abfolge des Auftretens verschiedener funktionaler Elemente im Spracherwerb, wie er beispielsweise im Englischen von Brown (1973)

beschrieben wurden, nicht Rechnung tragen. Gerkens Hypothese kann auch nicht erklären, warum die Artikelformen des Spanischen, die immer proklitisch, d. h. kein prosodisches Wort sind, nahezu ein halbes Jahr früher erworben werden, als die des Deutschen, die ein prosodisches Wort bilden können und damit prosodisch salienter sind (vgl. zum Spanischen und Deutschen Lleo & Demuth, 1999; zum Deutschen Penner & Weissenborn, 1996). Es scheinen also nicht allein prosodische Faktoren zu bestimmen, ob ein Funktionswort in einer kindersprachlichen Äußerung verwendet wird oder nicht.

Eine weitere Erklärungsmöglichkeit für die beobachtete Divergenz zwischen Sprachproduktion und Leistungen in rezeptiven Aufgabenstellungen könnte darin bestehen, dass der Produktion und der Rezeption von Sprache nicht das selbe syntaktische System zugrundeliegt. Diese Annahme erscheint zunächst weniger plausibel, da sie möglicherweise eine mehrfache und damit redundante Repräsentation syntaktischer Information beinhaltet. Dies ist jedoch kein Grund sie von vornherein abzulehnen. Auch im Bereich der Aphasie liegen Befunde vor, nach denen produktive und rezeptive syntaktische Leistungen bei erwachsenen Sprechern dissoziieren können (Miceli, Mazzucchi, Menn & Goodglass, 1983), die auf eine gewisse Unabhängigkeit der Systeme, die Produktion und Perzeption zugrundeliegen hindeuten. Für eine genaue Klärung dieser Unterschiede zwischen Produktion und Perzeption sind sicherlich weitere Forschungsarbeiten notwendig, es bleibt jedoch festzuhalten, dass nicht davon ausgegangen werden kann, dass Merkmale der spontanen Äußerungen von Kindern direkt das zugrundeliegende sprachliche Wissen reflektieren. Vielmehr kann dieses bereits viel komplexer sein als die Spontansprache erscheinen lässt (vgl. auch McKee & Iwasaki, 2001).

Fassen wir zunächst noch einmal die Hauptergebnisse aus den vorangehenden Kapiteln zusammen:

1. Sprachliche Äußerungen enthalten eine ganze Reihe von prosodischen sowie sequentiell-segmentalen Hinweisen, die prinzipiell für die Strukturierung der Äußerungen herangezogen werden können.
2. Kinder verfügen bereits in den ersten Lebensmonaten über erstaunliche Fähigkeiten, formale Aspekte ihrer Inputsprache zu analysieren, zu internalisieren und für den weiteren Spracherwerb zu nutzen. Wie unsere Untersuchungen zeigen, gilt dies auch für Funktionswörter, die weniger eine semantische, denn eine strukturelle Funktion haben.
3. Während der ersten Lebensmonate scheint die Informationsverarbeitung auf prosodische Merkmale des sprachlichen Inputs konzentriert. In der zweiten Hälfte des ersten Lebensjahres nimmt die Beachtung non-prosodischer sequentieller Information zu. In der zweiten Hälfte des ersten Lebensjahres ist ein zunehmendes Wissen über formale Merkmale der Zielsprache zu beobachten. Dieses Wissen fließt nun in die Sprachverarbeitungs- und Lernprozesse ein. Auch hier zeigt sich zuerst ein Einfluss prosodischen Wissens, das gegen Ende des ersten Lebensjahres von anderen Informationstypen ergänzt wird.
4. Der Vergleich zwischen den Verarbeitungs- und Erwerbsmechanismen im Kindes- und Erwachsenenalter zeigt weitgehende Parallelen. Alle Typen von Information, die Erwachsene für die basalen Segmentierungs- und Kategorisierungsschritte nutzen können, werden offensichtlich auch von Kindern bereits früh im Spracherwerb für denselben Zweck genutzt. Allerdings ist anzunehmen, dass bei erwachsenen Sprechern wissensgetriebene Mechanismen im Verhältnis zu datengetriebenen eine größere Relevanz besitzen. Bei Kindern muss zunächst von

einer Dominanz der datengetriebenen Mechanismen ausgegangen werden, die nach und nach durch wissensgetriebene ergänzt oder abgelöst werden. Dies geschieht offensichtlich kontinuierlich: sobald den Kindern sprachliches Wissen zur Verfügung steht, wird dies in den Analyseprozessen eingesetzt.

7.1 Früher Spracherwerb: Bestandsaufnahme und neue Perspektiven

7.1.1 Ein Modell des frühen Spracherwerbs

Bevor wir versuchen, uns ein generelles Bild über die Lern-, die Verarbeitungsmechanismen und die Wissensbasis zu machen, die den frühen Erwerb phonologischen und syntaktischen Wissens prägen, wollen wir zunächst die bisherigen Erkenntnisse über die sich während der ersten eineinhalb Lebensjahre in diesen Bereichen etablierenden und entfaltenden sprachlichen Fähigkeiten und Verarbeitungsprozeduren zusammenfassen (s. Tabelle 1).

Das Neugeborene ist mit einem auditiven System ausgestattet, das es ihm ermöglicht, die wesentlichen akustischen Merkmale, die in der Sprache zur Differenzierung zwischen verschiedenen Lauten, zwischen verschiedenen Betonungsmustern und für die Markierung struktureller Grenzen eingesetzt werden, wahrzunehmen und zu unterscheiden. Damit verfügt das Kind von Geburt an über die notwendigen perzeptuellen Basisfunktionen, die es ihm erlauben, seinen sprachlichen Input in Einheiten zu zergliedern und möglicherweise auch schon segmental spezifizierte mentale Repräsentationen für die extrahierten Einheiten zu bilden. Die Tatsache, dass Kinder schon früh auf ihren eigenen Namen reagieren, ist zwar bislang die einzige Evidenz dafür, dass Kinder schon in den ersten Lebensmonaten in der Lage sind, Langzeitrepräsentationen für häufig gehörte Einheiten aufzubauen, es ist aber zu vermuten, dass die ersten Langzeitrepräsentationen nicht unbedingt auf den eigenen Namen beschränkt sind. Es wurde vorgeschlagen, dass die frühe Repräsentation des eigenen Namens mit dessen Frequenz im kindlichen Input zu tun hat, so dass auch für andere hochfrequente Inputeinheiten sehr früh mentale Repräsentationen zu erwarten

Alter	Nachgewiesene Fähigkeiten
1. bis 4. Monat	<ul style="list-style-type: none"> - Diskrimination von rhythmischen Mustern - Diskrimination segmentaler Kontraste - Sensitivität für akustische Korrelate von strukturellen Grenzen - Erkennung prosodischer Merkmale der Muttersprache
5. Monat	<ul style="list-style-type: none"> - Erste Langzeitrepräsentationen hochfrequenter Inputsequenzen
7. Monat	<ul style="list-style-type: none"> - Erkennung der Grenzen von Intonationsphrasen - Extraktion und Repräsentation von Intonationsphrasen als Einheit - Rhythmische Segmentierung (einzelsprachunabhängig) - zielsprachliche Einflüsse auf Vokaldiskriminationsleistungen - Hörpräferenz für das in der Sprache dominierende Betonungsmster
8. Monat	<ul style="list-style-type: none"> - Wiedererkennung verschiedener Realisierungen einsilbiger Inhalts- und Funktionswörter in kontinuierlicher Sprache - Wiedererkennung trochäischer Zweisilbler in kontinuierlicher Sprache
9. Monat	<ul style="list-style-type: none"> - Nutzung non-prosodischer sequentieller Information für die Segmentierung und Strukturierung - Nutzung von Koartikulationsinformation für die Segmentierung und Strukturierung - Dominanz der Koartikulationsinformation gegenüber rein sequentieller Information
10. Monat	<ul style="list-style-type: none"> - Erkennung der Grenzen phonologischer Phrasen - Wissen über sprachspezifische phonotaktische Regularitäten - Nutzung phonotaktischen Wissens für Segmentierung - Rhythmische Segmentierung anhand des in der Zielsprache dominanten Betonungsmusters - Dominanz der rhythmischen Segmentierung über andere Segmentierungsmechanismen - Beginnende Integration prosodischer und sequentieller Segmentierungsinformation - Reduktion der Diskriminationsfähigkeit für nicht-native segmentale Konsonantenkontraste
12. Monat	<ul style="list-style-type: none"> - Dominanz der rhythmischen Segmentierung nicht mehr beobachtbar - Mentale – zumindest phonologische - Repräsentationen auch für Funktionswörter vorhanden - Nutzung der lexikalischer Repräsentationen für Segmentierung
16. Monat	<ul style="list-style-type: none"> - Syntaktisch-lexikalisches Wissen - Nutzung von Kookkurenzen zwischen Funktions- und Inhaltswörtern für die syntaktische Klassifizierung
18. Monat	<ul style="list-style-type: none"> - Erkennung von morpho-syntaktischen Dependenz- und Kongruenzrelationen

Tabelle 1: Perzeptive Fähigkeiten in den ersten 18 Lebensmonaten

sind. Es ist aber durchaus denkbar, dass auch andere Faktoren hier wirksam sind. So wäre denkbar, dass der eigene Name durch eine besondere Betonung, durch eine häufige isolierte Verwendung oder eine häufige äußerungsfinale Verwendung eine hohe perzeptuelle Salienz im Input erhält. Genauere Untersuchungen sind hier sicherlich von Nöten. In dieser frühen Phase zeigen sich auch bereits erste Hinweise darauf, dass die Kinder ihre eigene Muttersprache von anderen, auch prosodisch ähnlichen Sprachen differenzieren können, wobei sie wahrscheinlich bereits erworbenes Wissen über prosodische Merkmale ihrer Zielsprache heranziehen.

Sehr viel differenzierter ist das Bild der Fähigkeiten, die das 6 Monate alte Kind aufweist, vor allen Dingen in Bezug auf die Strukturierungsleistungen für den sprachlichen Input. In diesem Alter sind die Säuglinge in der Lage, die Grenzen von Intonationsphrasen zu erkennen, und sie scheinen Intonationsphrasen als Einheit zu repräsentieren. Da Intonationsphrasen im wesentlichen Sätzen oder Teilsätzen entsprechen, bilden die Repräsentationen dieser Einheiten auch syntaktischen Einheiten ab, so dass das Kind über Domänen verfügt, die für die Extraktion syntaktischer Regularitäten relevant sind. Auch auf kleinerer Ebene sehen wir die Nutzung von prosodischer Information für die Extraktion von Einheiten aus dem Input. So nutzt der Säugling den Wechsel zwischen betonten und unbetonten Silben für die Bildung komplexerer Einheiten: es werden bevorzugt solche Silbensequenzen als Einheit behandelt, die unterschiedlich stark betont sind und in einem konstanten Betonungsverhältnis zueinander stehen. Dabei spielt es beim 6 Monate alten Kind möglicherweise noch keine Rolle, ob die resultierende Einheit einen trochäischen oder einen jambischen Fuß bildet. Da vergleichbare Untersuchungen mit jüngeren Kindern fehlen, wissen wir nicht, ob die Fähigkeit zur rhythmischen Gliederung schon in den ersten Lebensmonaten vorhanden ist, oder ob sie sich erst um den 6. Lebensmonat herum etabliert. Wir sehen allerdings beim 6 Monate alten Kind bereits spezifisches prosodisches Wissen über Merkmale der Zielsprache: so erkennen 6 Monate alte Kinder bereits, welches Betonungsmuster in ihrer Zielsprache dominiert. Zudem zeigen sich erste Einflüsse des sich entwickelnden Phonemsystems auf die Wahrnehmung von lautlichen Unterschieden bei Vokalen.

In den nächsten beiden Lebensmonaten zeigt sich eine wachsende Sensitivität für nicht-prosodische sequentielle Information. So sind Säuglinge im Alter von 7

Monaten im Gegensatz zu 6-monatigen in der Lage, mentale Repräsentationen von Lautsequenzen schon nach einigen Hörwiederholungen aufzubauen und weitere Realisierungen dieser Sequenzen unter veränderten Bedingungen diesen Repräsentationen zuzuordnen. Zudem zeigt sich, dass Säuglinge diesen Alters rein sequentielle Information für die Extraktion von Einheiten aus kontinuierlichem sprachlichen Input nutzen können: so sind sie in der Lage, auch unbetonte Segmentfolgen als Einheiten zu erkennen und einen prosodisch vollkommen neutralen Sprachstrom anhand von Kookkurrenzmustern zu segmentieren. Allerdings dominiert in diesem Alter die Ausnutzung phonetisch-akustischer Merkmale des sprachlichen Inputs noch über statistisch segmentale Merkmale.

Eine Explosion des Erwerbs sprachspezifischen Wissens vor allen auf der segmentalen Ebene zeigt sich während des 9. Lebensmonats. Der 9 Monate alte Säugling kann native von nicht-nativen Lautsegmenten unterscheiden, der Einfluss des Phoneminventars der Muttersprache beginnt sich auch in der Klasse der Konsonanten mit der Abnahme der Diskriminationsfähigkeiten von nicht-nativen Konsonanten zu zeigen. Das Kind ist nun auch in der Lage, die Grenzen phonologischer Phrasen zu erkennen. Da diese recht gut mit syntaktischen Phrasen korrelieren, hat das Kind somit auch kleinere syntaktische Analyseeinheiten zur Verfügung. Das Kind hat Kenntnisse über die phonotaktischen Regularitäten der Zielsprache und über die Häufigkeit bestimmter Phonemabfolgen in bestimmten Wortpositionen erworben. Dieses Wissen wird für die Segmentierung des sprachlichen Inputs eingesetzt. Es zeichnet sich die beginnende Integration von prosodischen und nicht-prosodischen Informationsquellen ab, in Konfliktfällen vertraut das Kind jedoch noch stärker der prosodischen Information. Dabei hat sich die Segmentierung anhand des sprachspezifischen Betonungsmusters deutlich etabliert.

Im zwölften Lebensmonat ist diese Dominanz der Berücksichtigung prosodischer Information nicht mehr gegeben, was primär daran deutlich wird, dass sich keine Nachteile bei der Erkennung jambischer Wörter gegenüber trochäischen Wörtern mehr finden. Es zeigen sich Einflüsse des Wissens über die phonologische Form von Wörtern auf die Strukturierung des sprachlichen Inputs. In diesem Zusammenhang spielen Funktionswörter eine Rolle, die die Kinder im sprachlichen Input erkennen und somit als Ankerpunkte zu dessen Segmentierung und Strukturierung nutzen können. Es

ist anzunehmen, dass die frühe Fähigkeit, Funktionswörter im Input als bekannte Phonemsequenz zu erkennen, kein spezifisches Merkmal der Klasse von Funktionswörtern darstellt, sondern auf ihrer hohen Inputfrequenz und der damit verbundenen frühen mentalen Repräsentation beruht. Der gleiche Effekt ist für hochfrequente Inhaltswörter zu erwarten.

In der Zeit vom 16. bis zum 18. Lebensmonat zeigt sich eine deutliche Zunahme syntaktischen Wissens. Die Befunde zur Sprachperzeption machen hier vor allen Dingen deutlich, dass Kinder diesen Alters bereits über syntaktisch-lexikalisches Wissen verfügen sowie über Kenntnisse über Kookkurrenzen verschiedener Wortarten, d. h. Wissen über den Aufbau syntaktischer Phrasen in der Zielsprache. Dieses Wissen können sie nutzen, um neue Wörter, die in entsprechenden syntaktischen Kontexten verwendet werden, syntaktisch zu klassifizieren. Kurze Zeit später zeigen sich auch Sensitivitäten für syntaktische Abhängigkeiten zwischen verschiedenen Wortformen und Kongruenzbeziehungen. Auch dies belegt das zunehmende Wissen über die syntaktische Struktur der Zielsprache. Bislang ist jedoch das Bild der Veränderungen der sprachlichen Fähigkeiten in der ersten Hälfte des zweiten Lebensjahres sehr lückenhaft. Offensichtlich ist diese Lebensspanne aber eine wichtige Phase in der Etablierung ersten syntaktischen Wissens und der Ergänzung des bisher dominierenden phonologischen Wissens um weitere Komponenten.

Aus diesen einzelnen Befunden ergibt sich ein Bild der Entwicklung des phonologischen und syntaktischen Wissens und der Sprachverarbeitung während der ersten eineinhalb Lebensjahre, das durch drei verschiedene Phasen gekennzeichnet ist. Wir haben gesehen, dass die wesentlichen Perzeptionsmechanismen für die Wahrnehmung sprachlich relevanter akustischer Merkmale von Geburt an vorhanden sind. Aus den Untersuchungen zur Diskrimination sprachlicher Stimuli folgt auch, dass das Kind von Geburt an über die Mechanismen zum Aufbau kurzzeitiger mentaler Repräsentationen verfügen muss, denn die bei auditiver Darbietung notwendige sukzessive Präsentation der zu diskriminierenden Reize erfordert eine kurzfristige Gedächtnisspeicherung der Reize, über die der Vergleich vollzogen werden kann. Die Tatsache, dass bereits Neugeborene auf die Veränderung eines konstanten Merkmals in einer Kette von segmental verschiedenen Stimuli reagieren (z. B. Betonungsmuster und Silbenanzahl), belegt zudem, dass die Fähigkeit, einzelne Merkmale aus verschiedenen

sprachlichen Stimuli zu extrahieren und die Identität verschiedener sprachlicher Reize in Bezug auf diese Merkmale festzustellen, ebenfalls von Geburt an vorhanden ist. Innerhalb dieser Phase zeigen sich auch bereits die ersten Indikatoren des Aufbaus von Langzeitrepräsentationen, d. h. lexikalische Repräsentationen, für einzelne Wörter. Danach verfügen Säuglinge in diesem Alter zumindest über die Voraussetzungen und Mechanismen für den Aufbau des mentalen Lexikons. Die ebenfalls bereits im ersten Lebenshalbjahr vorhandene Sensitivität für prosodische Grenzmarkierungen und deren Nutzung für die Segmentierung des sprachlichen Inputs sorgen dafür, dass diese ersten mentalen Repräsentationen sprachstrukturell nicht vollkommen willkürliche Segmente darstellen. Gegen Ende dieser Phase sind zudem die ersten Anzeichen des Aufbaus sprachspezifischen Wissens im Bereich der Wortbetonungsmuster und im Phoneminventar bei Vokalen vorhanden. Abbildung 11 zeigt schematisch die sprachlichen Repräsentations- und Verarbeitungskomponenten, die das erste Lebenshalbjahr prägen.

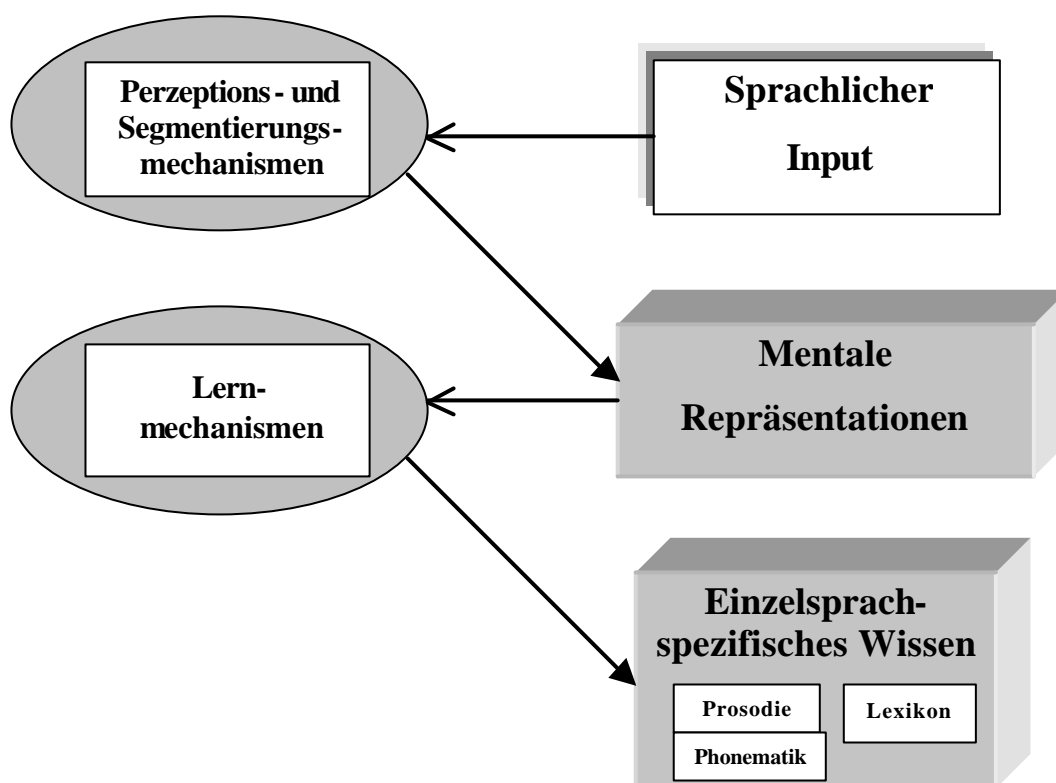


Abbildung 11: Sprachverarbeitung und Erwerb sprachspezifischen Wissens während der ersten 8 Lebensmonate

Die folgenden 6 Lebensmonate (vgl. Abbildung 12) sind geprägt von einer raschen Zunahme des einzelsprachlichen Wissens insbesondere im Bereich der Phonologie. Das phonematische Wissen erweitert sich insbesondere innerhalb des Konsonanteninventars, zudem werden phonotaktische Regularitäten aufgebaut und statistisches Wissen über die Häufigkeiten von Phonemkombinationen in bestimmten strukturellen Positionen erworben. Das mentale Lexikon umfasst gegen Ende der Phase auch bereits erste Repräsentationen von Funktionswörtern, die aufgrund ihrer relativ hohen Frequenz sogar einen beträchtlichen Anteil im Inventar des mentalen Lexikons darstellen sollten.

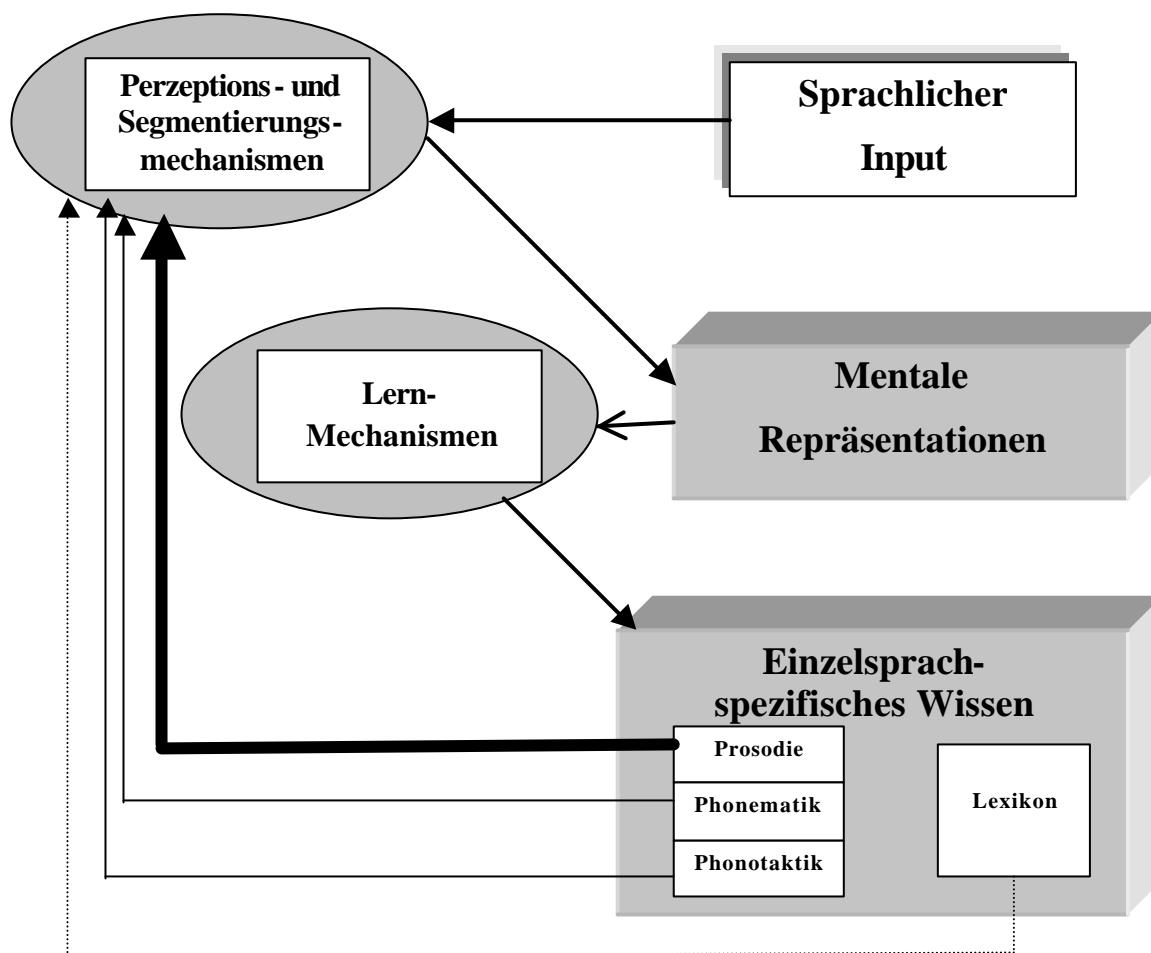


Abbildung 12: Sprachverarbeitung und Erwerb sprachspezifischen Wissens im zweiten Lebenshalbjahr

Die wesentlichste Veränderung gegenüber dem ersten Lebenshalbjahr besteht in der beginnenden Einflussnahme erworbenen sprachspezifischen Wissens auf die Verarbeitungsprozesse. Dies zeigt sich besonders deutlich für die Segmentierungsprozesse: schon zu Beginn des zweiten Lebenshalbjahres werden diese geleitet vom Wissen über das sprachspezifisch dominierende Betonungsmuster. In der zweiten Hälfte des zweiten Lebenshalbjahres ist der Erwerb phonotaktischen Wissens soweit vorangeschritten, dass auch dieses in die Segmentierungsprozesse einbezogen wird. Trotz beginnender Integration von prosodischen und nicht-prosodischen sequentiellen Strukturmerkmalen dominiert jedoch noch die Ausnutzung prosodischer Information.

Weitere bedeutende Veränderungen, die den Eintritt in eine neue Phase markieren, zeigen sich am Übergang vom ersten in das zweite Lebensjahr (vgl. Abbildung 13). Ein entscheidender Schritt vollzieht sich in Veränderungen in der Gewichtung der verschiedenen Informationsquellen, die zur Segmentierung herangezogen werden. Gegen Ende des ersten Lebensjahres verliert sich die vorherige Dominanz der Ausnutzung prosodischer Information gegenüber anderen Informationsquellen. So verfolgen die Kinder in diesem Alter keine primär rhythmisch geleitete Segmentierungsstrategie auf Wortebene mehr. Einer der Faktoren, die nun stärker in die Analyseprozesse eingehen, ist erstes phonologisch-lexikalisches Wissen, das zur Erkennung erster Wortformen in Input beiträgt und damit gleichzeitig Information über die Grenzen adjazenter Wortformen liefert. Bislang ist leider nichts darüber bekannt, welche anderen Informationstypen von den Kindern diesen Alters genutzt werden. Außerdem ist diese Phase geprägt durch den Erwerb syntaktischen Wissens. Bereits im ersten Drittel des zweiten Lebensjahres etabliert sich Wissen über die Wortklassenzugehörigkeit von lexikalischen Einträgen und über den Aufbau syntaktischer Phrasen. In der Mitte des zweiten Lebensjahres zeigt sich darüber hinaus zunehmende Sensitivität für morpho-syntaktische Abhängigkeits- und Kongruenzverhältnisse.

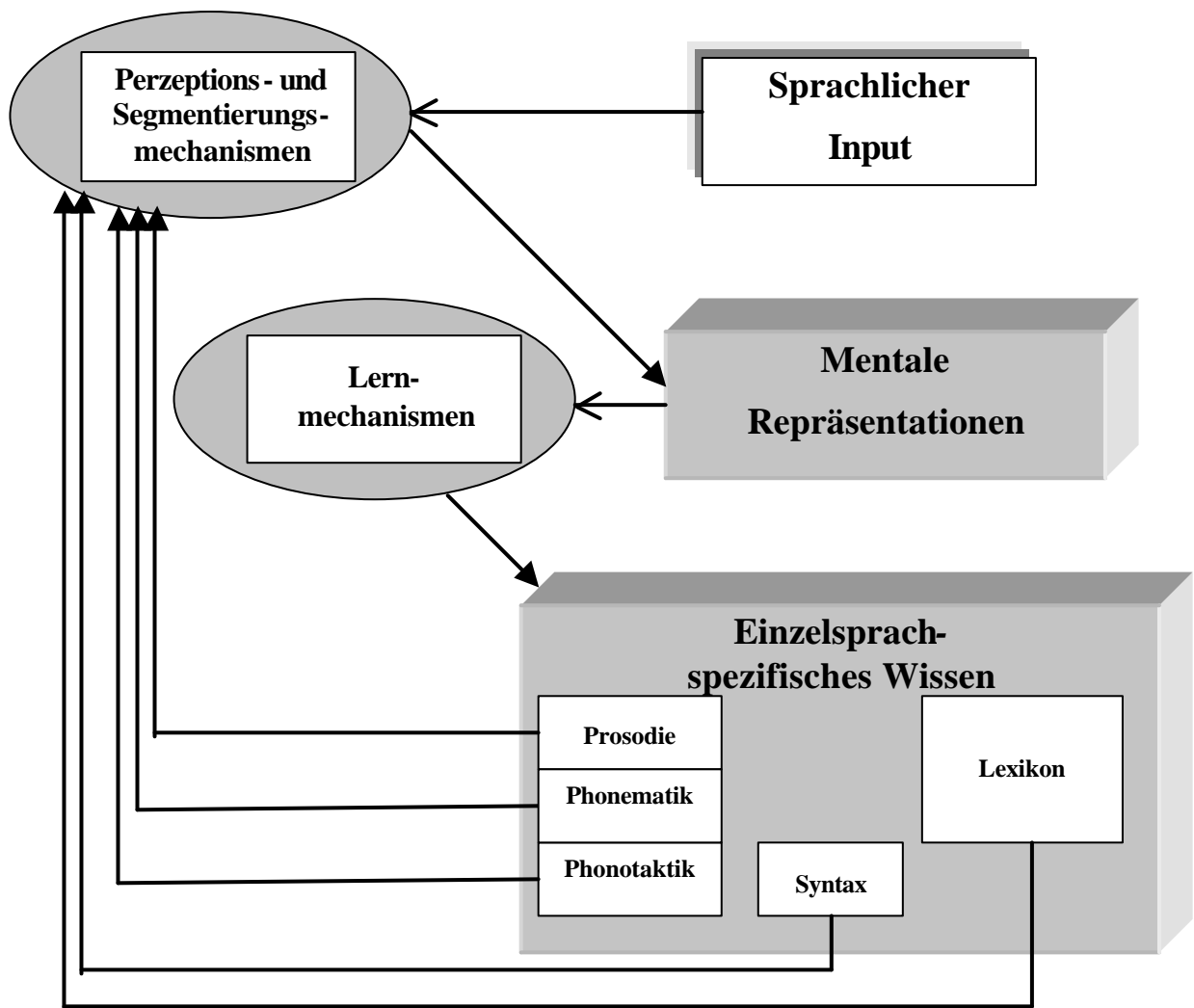


Abbildung 13: Sprachverarbeitung und Aufbau sprachspezifischen Wissens in der ersten Hälfte des zweiten Lebensjahres

7.1.2 Die Perzeptions- und Analysemechanismen

Nach dem oben skizzierten Modell ist der frühe Spracherwerb geprägt durch ein rasches Anwachsen sprachspezifischen Wissens, zunächst primär auf der Ebene der Phonologie, gegen Ende des ersten Lebensjahrs dann auch verstärkt in der Syntax. Ein weiteres Merkmal ist die Tatsache, dass das schon erworbene Wissen sobald es vorhanden ist, in die Verarbeitungs- und Analysemechanismen eingebunden wird und somit selbst den weiteren Aufbau des sprachlichen Systems vorantreibt.

Werfen wir nun einen genaueren Blick auf einige der Merkmale des vorgeschlagenen Modells. Eine wesentliche Aufgabe der weiteren Forschung wird es sicherlich sein, die angenommenen Perzeptions- und Lernmechanismen genauer zu spezifizieren. Den Lernmechanismen, zu denen neben distributionellen Analysemechanismen, auch Generalisierungs-, Induktions- und Analogiebildungsprozesse gehören, kommt ein zentraler Stellenwert in der Vermittlung zwischen den ersten, anhand der Perzeptionsmechanismen aufgebauten mentalen Repräsentationen und dem einzelsprachspezifischen Wissen zu. Vor dem Hintergrund, dass un gelenkte Lernmechanismen allein den Erwerb sprachlichen Wissens in verschiedenen Bereichen nicht erklären können (z. B. Gold, 1967; Gleitman, 1990; Markman, 1994), ist die Frage nach etwaigen Beschränkungen der Lernmechanismen von zentraler Bedeutung.

Für den Bereich von Phonologie und Syntax spielt in diesem Zusammenhang die Tatsache, dass der Input für die Lernmechanismen durch die vorhandenen Perzeptions- und Segmentierungsmechanismen bereits vorstrukturiert ist eine wesentliche Rolle. Nach unserem Modell sind die mentalen Repräsentationen, die als Input für die Lernmechanismen dienen, durch die Wirkungsweise der Segmentierungsmechanismen bereits Einheiten, die innerhalb des sprachlichen Systems auf verschiedenen Beschreibungsebenen relevant sind, d. h. sie entsprechen auch bereits in den frühesten Erwerbsphasen zumindest teilweise prosodischen, syntaktischen und auch lexikalischen Einheiten. Die Deckung zwischen den Einheiten der mentalen Repräsentation und strukturell relevanten Einheiten wird zudem durch den zunehmenden Einfluss bereits etablierten sprachlichen Wissens immer genauer, so dass den Lernprozessen gegen Ende des ersten Lebensjahres Repräsentationen zur Verfügung stehen, die wesentliche Strukturmerkmale der verschiedenen sprachlichen Ebenen enthalten und die vor allen Dingen Domänen darstellen, die für die Erfassung der Regularitäten auf den verschiedenen sprachlichen Ebenen relevant sind.

Zudem gibt es erste Hinweise darauf, dass distributionelle Analysemechanismen nicht jedes im Input vorhandene Distributionsmerkmal lernen. Dies gilt zum einen für sprecherspezifische Merkmale: so fanden Onishi, Chambers und Fisher (2002), dass Merkmale der Stimmqualität des Sprechers nicht in den Aufbau phonotaktischen Wissens einbezogen werden. Darüber hinaus berichten Newport und Aslin (2000) über

Befunde, nach denen Hörer nicht jede Art von statistischer Relation zwischen nicht adjazenten sprachlichen Elementen lernen. Nach diesen Ergebnissen werden dreisilbige Sequenzen nicht als Einheit repräsentiert, wenn zwischen der ersten und der dritten Silbe eine Übergangswahrscheinlichkeit von 100% besteht, aber die zweite Silbe variieren kann. Demgegenüber zeigte sich eine bessere Lernfähigkeit für die Verteilung nicht adjazenter Lautsegmente speziell von Konsonanten in Silbensequenzen mit variierenden Vokalen, beispielsweise für die konstante Abfolge der Konsonanten *d-k-b* in den Sequenzen *dowkiybae*, *daakuwebey* etc. Diese Verteilung entspricht dem morphologischen System des Hebräischen und des Arabischen. In diesen Sprachen werden morphologische Paradigmen anhand einer konsonantischen Wurzel gebildet, in die je nach Wortformen verschiedene Vokale präfigiert, infigiert oder suffigiert werden. Die Autoren vermuten, dass im Gegensatz zu den nicht adjazenten Lautsegmenten, nicht adjazente Silbenabfolgen in keiner Sprache ein systematisches strukturelles Merkmal darstellen. Wenn dies tatsächlich der Fall ist, würde sich in diesen Befunden eine hochinteressante Wechselwirkung zwischen möglichen sprachlichen Strukturmerkmalen und Beschränkungen der Lernmechanismen zeigen. Allerdings stammen die bisherigen Befunde, die auf Beschränkungen der Lernmechanismen hindeuten, ausschließlich von erwachsenen Sprechern, so dass nicht mit Sicherheit ausgeschlossen werden kann, dass diese Beschränkungen bereits ein Ausdruck der Vertrautheit mit bestimmten sprachlichen Merkmalen darstellen. Allerdings spielt im Englischen die Paradigmenbildung anhand konsonantischer Wurzeln keinerlei Rolle, so dass ein direkter Einfluss der erworbenen Sprache für die Befunde von Newport und Aslin, die ausschließlich Probanden mit englischer Muttersprache untersuchten, nicht allein ausschlaggebend sein können. Um festzustellen, ob diese Beschränkungen jedoch den Spracherwerb von vornherein lenken, ist die Überprüfung ihrer Wirksamkeit auch beim Säugling notwendig.

Neben diesen Beschränkungen innerhalb der Analysemechanismen selbst werden auch allgemein kognitive Beschränkungen der Gedächtnis- und Aufmerksamkeitsleistungen des Säuglings als relevant für den Erwerb morpho-syntaktischen Wissens diskutiert. Newport (1988, 1990) vermutet, dass beim Säugling eine Dissoziation zwischen hochgradigen Fähigkeiten zum Erwerb einer Sprache bei noch sehr eingeschränkten Fähigkeiten in vielen anderen kognitiven Domänen besteht,

d. h. beispielsweise generellen Gedächtnis- und Aufmerksamkeitsleistungen. Diese Beschränkungen führen zu einer Selektion bestimmter Ausschnitte aller vom Individuum wahrgenommenen Umweltreize und somit zu einer selektiven Verarbeitung bestimmter Aspekte des Inputs während andere Aspekte zunächst unbeachtet bleiben. Im Bereich der Morphologie unterstützen diese Kapazitätsbeschränkungen nach Newports Auffassung eine analytische Gliederung der Inputreize, die die Erkennung der morphologischen Struktur von Wörtern erleichtert. Damit wären die anfänglichen allgemein-kognitiven Beschränkungen also förderlich für den Erwerb gerade morpho-syntaktischen Wissens der Zielsprache.

Auch Elman (1993, 1999) ist der Überzeugung, dass die anfänglichen Beschränkungen in Gedächtnis- und Aufmerksamkeitsleistungen einen entscheidenden Einfluss auf den Verlauf des Spracherwerbs haben. Diese führen dazu, dass das Verarbeitungsfenster des Kindes sehr viel kleiner ist als das des Erwachsenen, das Kind seinen sprachlichen Input also in kleineren Einheiten verarbeitet als der erwachsene Sprecher. Diese kleineren Einheiten führen möglicherweise auf eine Fokussierung auf zunächst lokal bestehende Abhängigkeits- und Kongruenzverhältnisse, die bei Nachlassen der Verarbeitungsbeschränkungen dann auch in größeren Domänen etabliert werden können.

Tatsächlich zeigen sich bei Kindern Anzeichen relativ kleiner Verarbeitungsfenster. So fanden Santelmann und Jusczyk (1998) in einer Reihe von Untersuchungen, dass englische Kinder im Alter von 18 Monaten prinzipiell sensitiv für die Abhängigkeiten zwischen finitem und infinitem Verb sind, d. h. sie präferieren die korrekte Folge *is going* gegenüber einer falschen Folge wie *can going*. Allerdings zeigte sich diese Fähigkeit nicht mehr, wenn Sätze verwendet wurden, in denen ein mehr als zweisilbiges Adverb zwischen den beiden Verbformen auftrat. Dieser Befund deutet tatsächlich auf ein sehr kleines Verarbeitungsfenster von Kindern dieses Alters hin.

In diesen Zusammenhang gehört sicherlich auch die Frage, ob es Lernmechanismen gibt, die ausschließlich auf sprachlichen Input angewendet werden und somit domänenspezifisch sind. Die spärliche Evidenz in diesem Rahmen spricht bislang gegen eine Domänenspezifität. So zeigen verschiedene Untersuchungen, dass Säuglinge die gleiche prosodische und statistische Information, die sie für die Strukturierung sprachlicher Stimuli nutzen auch für die Strukturierung musikalischer

Reize heranziehen (Jusczyk & Krumhansl, 1993; Krumhansl & Jusczyk, 1990; Saffran, Johnson, Aslin & Newport, 1999). Darüber hinaus finden sich auch bei nicht menschlichen Spezies komplexe distributionsanalytische Fähigkeiten (Hauser, Newport & Aslin, 2001; Kluender, Lotto, Holt & Bloedel, 1998).

Weiterer Klärung bedarf vor allen Dingen auch die Frage, welche Aspekte des sprachlichen Wissens über diese Lernmechanismen erworben werden können. Nach den bisherigen Befunden ist es wahrscheinlich, dass distributionelles Lernen und Wissen in die Wortsegmentierung einfließt und in den Erwerb phonotaktischer Regularitäten. Neueren Untersuchungsbefunden zufolge, ist es wahrscheinlich, dass distributionelles Lernen auch beim Erwerb von Phonemkategorien eine Rolle spielt. So fanden Maye, Werker und Gerken (2002), dass bereits 6 Monate alte Säuglinge ein akustisches Kontinuum in zwei Kategorien teilen, wenn ihnen akustische Reize der Randbereiche des Kontinuums häufiger geboten werden als akustische Reize aus dem zentralen Bereich. Dagegen zeigten Säuglinge des gleichen Alters keine Anzeichen der Aufteilung des Kontinuums, wenn ihnen vornehmlich Reize aus dem zentralen Bereich geboten wurden. Im Bereich der Syntax haben wir gesehen, dass distributionelles Wissen in die Zuordnung von Wortformen zu syntaktischen Kategorien einfließt. Es ist wahrscheinlich, dass die distributionellen Analysemechanismen auch den Aufbau morphologischer Paradigmen, sowie den Erwerb von morphosyntaktischen Kongruenz- und Abhängigkeitsbeziehungen stützen. Der genaue Beitrag dieser Lernmechanismen zum Syntaxerwerb bedarf jedoch weiterer genauerer Untersuchungen. Diese werden auch mehr Klarheit darüber bringen, ob über diese gelenkten Lernmechanismen hinaus angeborenes strukturelles Wissen für den Erwerb des sprachspezifischen Regelsystems angenommen werden muss.

Zudem müssen wir davon ausgehen, dass wir bislang noch nicht alle möglichen Informationen aus dem Input, die den distributionellen Lernmechanismen zugänglich sind, kennen. Es sei daran erinnert, dass in der Modellierung der Lernbarkeit der Differenzierung zwischen Nomen und Verben von Durieux und Gillis (2001), die Klassifikation der Wörter am fehlerfreisten war, wenn dem Lernalgorithmus die gesamte, phonologische Form der Inputwörter zur Verfügung stand. Weniger erfolgreich gelang die Klassifizierung wenn der Lernalgorithmus ausschließlich Zugriff auf kodierte Information über beispielsweise Betonungsmuster, Wortlänge oder im

Wort auftretende Phonemklassen hatte. Dies bedeutet, dass der Lernalgorithmus dem Input Informationen entnehmen konnte, die über die kodierte Information hinausging. Welcher Art diese Information ist, ist bislang vollkommen unklar. Dieser Befund macht deutlich, dass weitere Studien zu Merkmalen des sprachlichen Inputs und ihrer möglichen Ausnutzbarkeit für Erwerbsmechanismen nötig sind.

7.1.2 Gewichtung und Integration von Hinweisreizen

Ein zweiter Aspekt des Modells, der weiterer Betrachtung bedarf, ist die Veränderung der Gewichtung bzw. der Beachtung der verschiedenen Informationstypen, die sich gegen Ende des ersten Lebensjahres vollzieht. Wie gesehen wird eine primäre Ausnutzung prosodischer Hinweisreize im zweiten Lebenshalbjahr durch eine stärkere Beachtung sequentieller Informationen abgelöst. Diesen Veränderungen in der Sprachverarbeitung wurde bislang am detailliertesten im sogenannten *Coalition-of-Cues*-Modell (Hirsh-Pasek & Golinkoff, 1996) Rechnung getragen. Diesem Modell der Entwicklung des Sprachverstehens liegt die Auffassung zugrunde, dass im Spracherwerb eine ganze Palette sprachlicher wie auch nonverbaler Hinweisreize in Form des sozialen und situativen Kontexts, in denen sprachliche Äußerungen getan werden, ausgenutzt werden. Veränderungen in der Beachtung dieser verschiedenen Hinweisreize erklären sich weder durch Veränderungen in den sprachlichen Äußerungen bzw. ihren nonverbalen Kontexten noch durch Veränderungen der Fähigkeit des Kindes, die verschiedenen Reize zu verarbeiten. Den Kindern stehen nach Auffassung der Autorinnen sämtliche Hinweisreize des sprachlichen und nicht-sprachlichen Inputs von Beginn an prinzipiell zur Verfügung, es ändert sich lediglich die Gewichtung, die diesen verschiedenen Hinweisreizen zukommt. Diese Veränderungen sind ein Ausdruck von Lernen, das nach einem prädeterminierten Schema verläuft, in dem Sinne, dass das Erwerbssystem in bestimmten Erwerbsphasen für bestimmte Typen von Information besonders sensitiv ist, wie es beispielsweise bei der Entwicklung des Vogelgesangs zu beobachten ist (Marler, 1991; Marler & Peters, 1987). So gehen Hirsh-Pasek und Golinkoff in ihrem Modell davon aus, dass die ersten neun Lebensmonate von einer besonders deutlichen Beachtung prosodischer Information gekennzeichnet sind, dieser Phase folgt eine stärkere Integration

semantischer Information in die Sprachverstehensprozesse. Erst danach, im Alter von ca. 24 Monaten vertraut das Kind verstärkt rein syntaktischer Information, die dann bei Widersprüchen der verschiedenen Informationstypen als die mit der höchsten Reliabilität behandelt wird.

Das von Hirsh-Pasek und Golinkoff für die Entwicklung des Sprachverstehens gezeichnete Bild der Ausnutzung der verschiedenen Typen von Information entspricht zumindest anfänglich dem hier vorgestellten Modell der Entwicklung der Sprachverarbeitung. Übereinstimmend wird angenommen, dass bis zu einem Alter von ca. 9 Monaten prosodische Information über andere Typen von Information dominiert. Für das Sprachverständnis gehen Hirsh-Pasek und Golinkoff jedoch dann von einer verstärkten Dominanz semantischer Information aus, und erst im dritten Lebensjahr vermuten sie eine stärkere Gewichtung syntaktischer Information im Verstehensprozess. Sicherlich ist davon auszugehen, dass ab dem Beginn des zweiten Lebensjahres auch das vorhandene und rapide wachsende semantisch-lexikalische Wissen (s. z. B. Fenson, Bates, Dale, Pethick, Reznick & Thal, 1994; Meints, Plunkett & Harris, 1999; Schafer, 1998) in die Sprachverarbeitungs- und Lernmechanismen integriert wird, aber nach unserem Modell wird für den Aufbau formal-sprachlichen Wissens in dieser Phase auch syntaktische und distributionell-segmentale Information intensiv genutzt. Mit dem Modell von Hirsh-Pasek und Golinkoff wird jedoch hier die Auffassung geteilt, dass die Kinder von Beginn an die Fähigkeit zur Verarbeitung dieser verschiedenen Informationstypen haben, die Veränderungen sich durch Verschiebungen der Gewichtungen ergeben, es also bestimmte Phasen gibt, in denen die Kinder für bestimmte Informationstypen besonders sensitiv sind. Dieser letzte Punkt und mögliche Ursachen für diese Veränderungen sollen hier weiter diskutiert werden.

Im Bereich der Sprachverarbeitung haben wir die Veränderungen der Gewichtung verschiedener Hinweisreize besonders deutlich im Bereich der Segmentierung des sprachlichen Inputs gesehen. So verlassen sich Kinder im Alter von ca. 8 Monaten bei der Segmentierung kontinuierlicher Sprache primär auf prosodische Hinweise, selbst wenn diese – gemessen an den standardsprachlichen lexikalischen Formen – zu Fehlsegmentationen führen. Dass diese Segmentierungsmechanismen ohne Rückgriff auf anderes einzelsprachspezifisches Wissen angewendet werden können, belegen unsere Befunde, nach denen auch amerikanische Säuglinge in der Lage sind,

aus deutschsprachigen Texten trochäische Wörter zu segmentieren. Da keins der Kinder vorher Kontakt zu Sprechern des Deutschen hatte, ist auszuschließen, dass in diesem Alter die Segmentierung des sprachlichen Inputs bereits erste wissensgetriebene Analyseprozesse voraussetzt. Wie gezeigt, ist im Alter von 12 Monaten das primäre Vertrauen in prosodische Information verschwunden. Es hat eine Umorientierung zu einer stärkeren Beachtung anderer Hinweisreize stattgefunden.

Wie lässt sich die anfängliche Dominanz der prosodischen Information und die erst spätere Verwendung nicht-prosodischer distributioneller Information erklären? Der zeitliche Unterschied in der Ausnutzung prosodischer und segmental-sequentieller Information könnte damit in Verbindung stehen, wie schnell in diesen beiden Bereichen zielsprachliches Wissen aufgebaut werden kann. Dass dies für prosodische Information schneller geschieht als für nicht-prosodische, hängt wahrscheinlich zum einen mit der früheren und direkteren Wahrnehmbarkeit der akustisch-prosodischen Merkmale zusammen. Wie in Kapitel 4 angesprochen, kann man davon ausgehen, dass das Kind bereits pränatal, sobald sein akustisches System ausreichend ausgereift ist, dazu in der Lage ist, sprachliche Äußerungen und vor allen Dingen deren prosodische Information, wahrzunehmen. Damit hat das Kind bei seiner Geburt wahrscheinlich bereits ca. 2 bis 3 Monate Erfahrungen mit den prosodischen Merkmalen seiner Muttersprache gesammelt. Segmentale Eigenschaften sprachlicher Äußerungen sind intra-uterin dagegen nicht gut wahrnehmbar, so dass der Erwerb segmentaler Eigenschaften der Zielsprache wahrscheinlich erst mit der Geburt beginnen kann. Dieser Zeitraum von zwei bis drei Monaten deckt sich ziemlich genau mit der Verzögerung, den wir im Erwerb prosodischer Eigenschaften der Zielsprache gegenüber sequentiell-segmentalen beobachtet haben, denn nach unseren Befunden haben Säuglinge im Alter von 6 Monaten Wissen über das in ihrer Sprache dominierende Betonungsmuster während Wissen über phonotaktische Regularitäten erst im Alter von 9 Monaten nachgewiesen wurden.

Zweitens erfordert der Aufbau sequentiell-segmentalen Wissens in der beschriebenen Form vermutlich feinere Analyseprozeduren und Repräsentationsmechanismen als der Aufbau von Wissen über dominante Betonungsstrukturen. Dies liegt zum einen daran, dass Betonungsinformation aufgrund ihrer Verbindung mit konkreten akustischen Parametern dem Input direkter zu

entnehmen ist als Wissen über die Häufigkeit einzelner Segmente oder von Segmentkombinationen, die nur über den Vergleich einer größeren Menge von sprachlichen Einheiten bzw. deren mentalen Repräsentationen zu ermitteln sind, d. h. der Aufbau dieses Wissens setzt bereits ein größeres Repertoire an mentalen Formrepräsentationen voraus. Zudem ist die Zahl der möglichen Betonungsmuster in einer Sprache im Vergleich zur Zahl der möglichen Phonemabfolgen sehr eingeschränkt. So werden in der Phonologie nicht mehr als drei Akzentwerte angenommen: eine Silbe kann entweder den Primärakzent in einem Wort tragen, sie kann einen Sekundärakzent tragen oder sie ist unbetont. Die Zahl der möglichen Betonungsmuster einer Silbenabfolge ist also sehr eingeschränkt, so dass sich ein dominierendes Betonungsmuster bereits anhand eines sehr viel geringeren sprachlichen Inputs feststellen lässt, als dies für Kombinationsmöglichkeiten auf segmentaler Ebene gilt, die allein aufgrund des größeren Inventars an Einheiten weitaus vielfältiger sind.

Eine unterschiedliche Salienz der verschiedenen Informationstypen im sprachlichen Input in verschiedenen Phasen des Erwerbs vermuten Werker, Lloyd, Pegg und Polka (1996) als Ursache für die Verschiebung der Wertung der Hinweisreize. Sie gehen davon aus, dass die globalen prosodischen Merkmale der Sprache durch ihre Auffälligkeit und ihre direkte akustische Verfügbarkeit zunächst die Verarbeitungsprozesse des Säuglings dominieren. Wenn der Säugling in diesem Bereich die Merkmale seiner Zielsprache erworben hat, wird detaillierte Information innerhalb dieser Strukturen salienter und den Erwerbsprozeduren zugänglich. Auch hier werden Veränderungen in den Erwerbs- und Verarbeitungsprozessen nicht durch Veränderungen in den perzeptuellen Fähigkeiten erklärt, sondern durch Veränderungen in der Fokussierung einzelner Aspekte des sprachlichen Inputs, die auch durch das bereits erworbene Wissen bedingt sind.

Der Übergang von der Dominanz eines Typs von Hinweisreizen zu einem anderen ist nicht abrupt sondern eher fließend. So verlassen sich die Kinder – selbst wenn ihnen segmental-distributionelle Analysemechanismen und entsprechendes Wissen schon zur Verfügung stehen, doch zunächst stärker auf die „altvertrauten“ akustisch direkt zugänglichen Merkmale. So zeigte sich in der Untersuchung von Mattys et al. (1999), dass 9 Monate alte Kinder sowohl prosodische Information zur Wortsegmentierung heranziehen als auch Wissen darüber, dass bestimmte

Phonemkombinationen häufiger wortintern oder wortübergreifend vorkommen. Wurden diese beiden Typen von Information im Stimulusmaterial in widersprüchlicher Weise verwendet, vertrauten die Kinder stärker der prosodischen Information. Ein ähnliches Muster fand sich in der Untersuchung von Johnson und Jusczyk (2001), die Koartikulationsinformation mit statistischer Abfolgeinformation in Widerspruch setzten. Es zeigte sich unter dieser Bedingung, dass sich die 8 Monate alten Säuglinge stärker an der Koartikulationsinformation orientierten, d. h. Silbensequenzen mit natürlichen Koartikulationsmerkmalen als stärkere Einheit betrachteten, als Silbensequenzen mit höherer interner Übergangswahrscheinlichkeit aber ohne Koartikulationsmerkmale.

Ein entscheidender Schritt für die vollständige Ausnutzung der Input-Information besteht jedoch darin, die Information verschiedener Typen von Hinweisreizen integrieren und gleichzeitig ausnutzen zu können, denn dies führt zu einer weitaus reliableren Analyse des sprachlichen Inputs als das Vertrauen auf nur eine Informationsquelle. So fanden Christiansen, Allen und Seidenberg (1998) in ihrer bereits erwähnten Studie zu Segmentierungsprozessen an einem Korpus kindgerichteter Sprache deutliche Interaktionen verschiedener Informationstypen, die der Trainingskorpus enthielt. Ihr Algorithmus erzielte die besten Wortsegmentierungsleistungen wenn das Trainingskorpus eine phonologische Kodierung jedes Lautsegments sowie Betonungsinformation enthielt. Bei Trainingskorpora, die jeweils nur einen dieser Informationstypen enthielt, waren die Segmentierungsleistungen weitaus schlechter. Interessanterweise lagen die Leistungen für das Korpus mit beiden Informationstypen höher, als die Ausnutzung der einzelnen Hinweistypen in den Korpora mit nur einem Informationstyp erwarten ließen. Dies bedeutet, dass die gleichzeitige Ausnutzung verschiedener Informationsquellen nicht nur rein additiv zu einer genaueren Analyseleistung führt, sondern dass die verschiedenen Informationsquellen interaktiv in die Analyseprozeduren integriert werden können und zu besseren Analyseleistungen führen als jeder Typ von Hinweisreiz für sich genommen.

Ein ähnliches Bild zeichnet sich in der Modellierung der Wortkategorisierung ab. Auch Durieux und Gillis (2001) berichten über erheblich bessere korrekte Wortkategorisierungen, wenn dem Algorithmus mehrere Hinweise auf die Wortklasse

zur Verfügung stehen. So stellten sie fest, dass der Algorithmus weitaus besser in der Lage war, zu entscheiden, ob ein englisches Wort ein Nomen oder ein Verb ist, wenn ihm das Betonungsmuster, die Vokalhöhe, die Qualität der Konsonanten und die Anzahl der Segmente als Information zur Verfügung stand, als bei jedem dieser Informationstypen in Isolation. Bei den Verben veränderte sich die Anzahl der korrekten Zuordnungen bei der Kombination aller Informationstypen sogar dramatisch. War bei Berücksichtigung nur eines einzigen Informationstyps in keiner Bedingung der Anteil der korrekt identifizierten Verben über 35%, so lag dieser Anteil bei der Kombination aller Quellen bei 61%. Bei der Nomenidentifikation zeigte sich dieser erhebliche Vorteil der Informationskombination dagegen nicht.

In Bezug auf den Spracherwerb wissen wir bislang kaum etwas darüber, wann und auf welche Weise Säuglinge beginnen, Hinweise aus verschiedenen Informationsquellen in ihren Verarbeitungsprozessen zu integrieren. Einzig in der Studie von Morgan und Saffran (1995) zeigen sich erste Ergebnisse, die frühe Integrationsfähigkeiten andeuten. Die Forscher fanden, dass 9 Monate alte Säuglinge Zweisilber nur dann als Einheit repräsentieren, wenn sowohl prosodische als auch segmental-sequentielle Information diese Analyse unterstützen. Demgegenüber zeigten 6 Monate alte Säuglinge noch keine Anzeichen dafür, dass sie diese verschiedenen Informationstypen gleichzeitig beachten. Die Integrationsfähigkeit für verschiedene Hinweisreize ist sicherlich eine wichtige Voraussetzung dafür, dass ein Typ von Information seine Dominanz gegenüber anderen verliert und somit Wege für den weiterführenden Erwerb von strukturellen Eigenschaften der Zielsprache geöffnet werden. Welche genauen Faktoren zu diesen Veränderungen in der Beachtung und Integration von Eigenschaften des sprachlichen Inputs beitragen und wie sich diese Veränderungen genau vollziehen, ist bislang weitgehend unbekannt. Eine Klärung dieser Faktoren wird sicherlich einer der zentralen Gegenstände der frühen Spracherwerbsforschung der nächsten Zukunft sein.

In der Verschiebung der Beachtung von Hinweisreizen ist sicherlich auch die Reliabilität der einzelnen Informationstypen innerhalb einer gegebenen Sprache von Bedeutung. Warum sollte ein Kind, in dessen Zielsprache das Wortbetonungsmuster ein sehr zuverlässiger Hinweis auf Wortgrenzen darstellt, das Vertrauen in diesen Informationstyp zugunsten eines weniger reliableren aufgeben? Es ist zu vermuten, dass

Veränderungen in den Dominanzen auch strukturelle Eigenschaften der Zielsprache widerspiegeln, d. h. dass das Verarbeitungssystem seine Mechanismen den strukturellen Erfordernissen und Gegebenheiten der Zielsprache anpasst. Diesem Aspekt wird im folgenden Abschnitt weiter nachgegangen.

7.1.3 Der Einfluss einzelsprachspezifischen Wissens auf die Sprachverarbeitung und den weiteren Spracherwerb

Die Adaptation der Sprachverarbeitung an zielsprachliche Gegebenheiten wurde bislang hauptsächlich im Bereich der Verarbeitung phonologischer Information evident. In zahlreichen Studien zeigte sich, dass erwachsene Sprecher bestimmte nicht-native phonemische Kontraste schlechter diskriminieren können als native Kontraste (als Überblick s. z. B. Werker, 1991), während sich bei Neugeborenen noch universelle Diskriminationsfähigkeiten finden (als Überblick s. z. B. Jusczyk, 1997), d. h. Säuglinge sind in der Lage, weitaus mehr lautliche Differenzierungen zu unterscheiden als erwachsene Sprecher. Wie vor allen Dingen die Arbeiten von Werker und Mitarbeitern (z. B. Werker & Lalonde, 1988; Werker & Tees, 1984) gezeigt haben, finden die kritischen Veränderungen, d. h. ein Nachlassen der Diskriminationsfähigkeiten für muttersprachlich nicht relevante Kontraste, bereits im ersten Lebensjahr statt. Die gleichen nicht-nativen Konsonantenkontraste, die sechs Monate alte Kinder noch ohne Probleme unterscheiden konnten, werden von 12 Monate alten Kindern nicht mehr diskriminiert. Für Vokale etabliert sich der Einfluss des phonologischen Systems der Zielsprache auf die Wahrnehmung noch wesentlich früher. Wie schon in Kapitel 6 angesprochen, ist die Abnahme dieser Diskriminationsfähigkeit aber nicht für alle segmentalen Kontraste zu beobachten. Für die Abnahme bzw. den Erhalt der Diskriminationsfähigkeit ist offensichtlich die Nähe eines Lautkontrasts zum phonologischen System der Muttersprache bestimmend (Best, 1990; Best, McRoberts & Sithole, 1988), d. h. je deutlicher ein bestimmter Kontrast sich vom muttersprachlichen Lautinventar abhebt, desto einfacher ist er auch für Erwachsene zu diskriminieren. Die Diskriminationsfähigkeiten sind also über das phonologische System des Sprachbenutzers vermittelt. Der Hörer versucht, jeden Laut, den er wahrnimmt, auf sein phonologisches System abzubilden. Gelingt dies nicht, da der gehörte Laut zu weit vom

eigenen System entfernt liegt, beispielsweise für einen deutschen Sprecher Klicklaute, zeigen sich gute Diskriminationsleistungen. Können zwei zu diskriminierende Laute jedoch auf das eigene System abgebildet werden und fallen dort in eine Kategorie, werden sie nicht diskriminiert.

Wie in Kapitel 3 dargestellt finden sich einzelsprachliche Adaptationen an strukturelle Gegebenheiten der Zielsprache auch im Bereich der Segmentierungsmechanismen, die beim erwachsenen Sprecher ebenfalls durch das phonologische System der Muttersprache beeinflusst sind. Wie gezeigt nutzen erwachsene Sprecher die Einheiten für die Segmentierung, die in ihrer Sprache rhythmisch relevant sind. Cutler und Mitarbeiter (Cutler et al., 1992) vermuten, dass sich die Segmentierungsmechanismen in Form einer Parameterfixierung während des frühen Spracherwerbs an die sprachspezifischen phonologischen – insbesondere rhythmischen Strukturen - der Zielsprache adaptieren. Für diese Hypothese spricht die beobachtete Inflexibilität der Segmentierungsmechanismen bei erwachsenen Sprechern: so wenden erwachsene Sprecher ihre Segmentierungsstrategie auch auf Sprachmaterial an, dessen phonologische Struktur für diese Segmentierungsroutinen nicht optimal ist (Cutler et al., 1986). Direkte empirische Evidenz für die Hypothese einer Fixierung des Segmentierungsmechanismus während des frühen Spracherwerbs fehlt bislang, allerdings scheinen Neugeborene - wie gesehen - tatsächlich besonders sensitiv gegenüber rhythmischen Eigenschaften der Sprache zu sein, denn sie sind in der Lage, rhythmisch unterschiedliche Sprachen zu differenzieren, nicht aber rhythmisch identische (Nazzi et al., 1998).

Auch auf morpho-syntaktischer Ebene finden sich Hinweise auf Einflüsse sprachspezifischer Merkmale auf Verarbeitungsmechanismen. Zahlreiche Studien, die im Rahmen des *Competition-of-Cue* Modells (s. Bates & MacWhinney, 1989) durchgeführt wurden, zeigen sprachspezifische Unterschiede in der Ausnutzung verschiedener Informationen (Wortfolge, Agreement, Kasus) bei der Satzinterpretation, speziell bei der Zuweisung thematischer Rollen, die offensichtlich von syntaktischen und morpho-syntaktischen Merkmalen der Muttersprache des Hörers bedingt sind. Die gefundene zwischensprachliche Varianz hängt damit zusammen, wie zuverlässig ein gegebener Hinweis in der jeweiligen Sprache für die Satzinterpretation ist. Beispielsweise verlassen sich englische Sprecher bei der Subjektidentifikation strikt auf

die Wortfolge, indem sie eine präverbale Nominalphrase als Subjekt interpretieren, während Hörer von Sprachen mit stärkerer Kasusmorphologie oder stärkerer Subjekt-Verb-Kongruenz eher den flexionsmorphologischen Markierungen folgen. Entsprechende Unterschiede zwischen Sprechern verschiedener Sprachen wurden bereits bei Kindern ab 2 bis 3 Jahren festgestellt (s. den Überblick in Bates & MacWhinney, 1989). Bates und McWhinney vermuten, dass die Erwerbsreihenfolge in der Ausnutzung dieser Hinweisreize und der Grad ihrer Ausnutzung durch den erwachsenen Sprecher mit ihrer Reliabilität und der Deutlichkeit ihrer Markierung (Wahrnehmbarkeit) in der jeweiligen Zielsprache zu tun hat. Es wäre aber auch denkbar, dass die schon bei Kindern gefundenen Unterschiede in den Interpretationsstrategien nicht den Erwerb der unterschiedlichen Markierungsmöglichkeiten reflektieren, sondern eine "Desensibilisierung" für Hinweisreize, die in der jeweiligen Zielsprache nicht bzw. wenig vorkommen oder wenig reliabel sind. Nach dieser Hypothese wären anfänglich keine Unterschiede zwischen Kindern verschiedener Zielsprachen in Bezug auf die Sensitivität für unterschiedliche Informationstypen (Reihenfolge, Flexionsendungen etc.) zu erwarten, sondern diese sollte sich erst mit zunehmendem Einfluss des sprachlichen Inputs im Verlauf des frühen Spracherwerbs zeigen.

Alle diese Befunde deuten darauf hin, dass strukturelle Merkmale der Zielsprache die Verarbeitungsmechanismen beeinflussen, indem sich diese den Merkmalen der Zielsprache anpassen und so ihre Wirkungsweise optimieren. Diese Interaktion zwischen Merkmalen des Lerngegenstandes und Merkmalen des Erwerbssystems ist bislang in der Forschung viel zu wenig beachtet worden und erfordert sprachvergleichende Untersuchungen zum Erwerb von strukturell divergierenden Sprachen.

In diesem Zusammenhang wäre natürlich auch die Untersuchung des bilingualen Spracherwerbs von erheblichem Interesse. Unter der Annahme einer frühen Einflussnahme sprachspezifischer Eigenschaften auf den Erwerbsprozess wäre für den bilingualen Erwerb eine möglichst frühzeitige Separation der beiden sprachlichen Systeme vorteilhaft (vgl. Mehler et al., 1996). Dies sollte dem Säugling nach unserem Modell am leichtesten fallen, wenn seine beiden Muttersprachen sich bereits in ihren globalen prosodischen Merkmalen deutlich unterscheiden. Für diesen Fall könnte sich

eine erfolgreiche Differenzierung der beiden Sprachen bereits in den ersten Lebensmonaten vollziehen, so dass von Anfang an separates Wissen für die beiden Zielsprachen aufgebaut wird. Größere Probleme sollten sich dagegen bei prosodisch sehr ähnlichen Sprachen finden, die sich jedoch in anderen strukturellen Bereichen deutlich unterscheiden. Hier wird möglicherweise während der prosodisch dominierten Phase ein einziges System für beide Sprachen aufgebaut. Der Erwerb struktureller Unterschiede zwischen den beiden Sprachen sollte dadurch erschwert werden.

Untersuchungen zu den frühesten Phasen des bilingualen Spracherwerbs deuten allerdings darauf hin, dass bilingual aufwachsende Kinder auch dann bereits in den ersten Lebensmonaten die beiden Zielsprachen voneinander diskriminieren können, wenn diese in ihren globalen prosodischen Merkmalen übereinstimmen (Bosch & Sebastián-Gallés, 2001). Allerdings zeigten die bilingualen Kinder nicht die für ihre monolingualen Altersgenossen typische, schnellere Orientierung zu Reizen der Muttersprache gegenüber Reizen einer Fremdsprache. Dies könnte darauf hindeuten, dass das viermonatige bilinguale Kind noch keine so deutliche Präferenz für seine Muttersprachen gegenüber einer anderen Sprache entwickelt hat wie das monolingual aufwachsende. Dies ist ein erster Indikator für einen frühen Unterschied in der Sprachperzeption bei bilingualen gegenüber monolingualen Kindern.

Interessant wäre in diesem Zusammenhang natürlich eine genauere Untersuchung des Verlaufs des bilingualen Spracherwerbs bei Kindern mit unterschiedlichen Kombinationen von Inputsprachen. Nach unserem Modell wären deutliche Unterschiede im Verlauf des bilingualen Spracherwerbs in Abhängigkeit von den spezifischen strukturellen Gemeinsamkeiten und Unterschieden der beiden Zielsprachen zu erwarten. Hier ließen Untersuchungen des bilingualen Spracherwerbs sicherlich auch generelle Rückschlüsse darüber zu, inwieweit der Aufbau des sprachlichen Systems von Merkmalen des sprachlichen Inputs beeinflusst wird und wie flexibel die Interaktion zwischen Verarbeitungsprozeduren und diesen Inputmerkmalen sein kann.

7.2. Eine kritische Periode für den Spracherwerb?

Eine Vielzahl von Befunden weist darauf hin, dass die Fähigkeit, eine Sprache zu erwerben, im Erwachsenenalter nicht mehr in der Weise gegeben ist, wie sie beim Kind zu beobachten ist. Besonders im morpho-syntaktischen Bereich führt ein später Beginn des Spracherwerbs häufig zu einer geringeren sprachlichen Kompetenz als dies für den normal im Säuglingsalter einsetzenden Spracherwerb gilt. Dies zeigt sich beim erst späten Einsatz des Erwerbs einer gesprochenen Sprache, wie er in Einzelfällen bei Menschen, die isoliert ohne sprachlichen Input aufgewachsen waren, beobachtet werden konnte (Curtiss, 1977), in Unterschieden im Erwerbserfolg für die Gebärdensprache Taubstummer in Abhängigkeit vom Alter, in dem die Lerner ersten Gebärdeneinput erhielten (Newport 1988; Newport, 1991), sowie beim Zweitspracherwerb (z. B. Bialystok & Hakuta, 1999; Johnson & Newport, 1989; Birdsong & Mollis, 2001).

Die Altersabhängigkeit des Erfolgs im Aufbau sprachlicher Kompetenz wurde vielfach im Sinne einer kritischen Periode für den Spracherwerb interpretiert, für deren Grenzen verschiedenste Faktoren wie Veränderungen in der neuronalen Plastizität, Veränderungen in den Lernmechanismen und Interferenzen durch bereits vorhandenes sprachliches und nicht-sprachliches Wissen beim späteren Erwerb diskutiert wurden (z. B. Lenneberg, 1967; als Überblick s. Birdsong, 1999). Mittlerweile mehren sich allerdings Befunde, nach denen auch Zweitspracherwerber, syntaktische und morpho-syntaktische Kompetenzen in der Zweitsprache erreichen können, die selbst bei einer gezielten Testung dieser Fähigkeiten nicht von der Kompetenz eines Muttersprachlers zu differenzieren sind (als Überblick s. Birdsong, 1999). Aktuelle Ergebnisse neurophysiologischer Untersuchungen deuten zudem darauf hin, dass erwachsene Lerner die gleichen Verarbeitungsmuster für morpho-syntaktische Information zeigen können, wie sie charakteristisch für die Verarbeitung entsprechender Informationen in der Muttersprache sind (Friederici, Steinhauer & Pfeifer, 2002; Hahne, 2001). Diese Beobachtungen machen das strikte Konzept einer kritischen Periode für den Spracherwerb, nach der aus biologisch determinierten Gründen der Reifung des Systems ein voller Spracherwerb nicht mehr möglich ist, fragwürdig.

Der Vergleich, der in den vorherigen Kapiteln in Bezug auf die Verarbeitungs- und auch Lernkapazitäten zwischen erwachsenen Sprechern und Säuglingen gezogen wurde, bringt jedoch Ergebnisse, die die Erklärungsmöglichkeiten für die sich im

allgemeinen abschwächende Fähigkeit, formale Aspekte einer Sprache zu erwerben einzuschränken helfen. Wie beispielsweise Newport (1991) betont, könnten Veränderungen der Lernfähigkeit für bestimmte Domänen aus einer Veränderung der kritischen Lernmechanismen selbst resultieren oder aus Veränderungen anderer kognitiver Kapazitäten, die in irgendeiner Weise auf den Lernerfolg Einfluss nehmen. Dies ist der Kernpunkt ihrer bereits in Abschnitt 7.1.1 diskutierten Auffassung, nach der beispielsweise generelle Beschränkungen der Gedächtniskapazität eine anfängliche Analyse kleinerer Einheiten im kindlichen Sprachverarbeitungsmechanismus gewährleisten, was gerade für den Erwerb morpho-syntaktischer Fähigkeiten förderlich sein könnte. Ein Wegfall dieser allgemeinen Beschränkungen hätte damit Nachteile für den sich nicht analog verändernden Erwerbsmechanismus.

Das Bild, das sich aus unserem Vergleich zwischen erwachsenem und kindlichem Sprachverarbeiter und -lerner ergibt, spricht prinzipiell für Newports Annahmen. Wir haben gesehen, dass sich Kinder und Erwachsene nicht grundsätzlich in der Art der Information, die sie für die strukturelle Analyse des sprachlichen Inputs nutzen können, unterscheiden. Ein Unterschied scheint lediglich das Verhältnis von wissens- und datengetriebenen Prozessen für die Verarbeitung zu sein: dieses ist beim sehr jungen Kind sicherlich deutlich zugunsten datengetriebener Prozesse anders strukturiert als beim älteren Kind und beim Erwachsenen. Auch scheint der erwachsene Sprecher über ähnliche distributionsanalytische Fähigkeiten zu verfügen. Wie die in Abschnitt 5.1.4 dargestellten Untersuchungsbefunde gezeigt haben, ist auch der erwachsene Sprecher in der Lage, aus nur einer relativ geringen Menge sprachlichen Inputs formale Aspekte der zugrundeliegenden Grammatik zu erkennen. Dabei scheint der Erwachsene ebenso wie der Säugling distributionelle Merkmale von Kookkurrenzen und Häufigkeiten nutzen zu können. Beide Aspekte unterstreichen, dass nicht von einer grundsätzlichen Veränderung der Verarbeitungs- und Lernmechanismen, die zum Spracherwerb beitragen vom Säuglings- zum Erwachsenenalter ausgegangen werden kann. Dieses macht die Annahme, dass eher Veränderungen in den Randbedingungen für die Verarbeitungs- und Lernmechanismen für eine eventuelle Abnahme der Lernfähigkeit relevant sein könnten.

Neben den von Newport bereits diskutierten Einschränkungen könnte ein weiterer wichtiger Unterschied zwischen dem Säugling und einem älteren Sprachlerner

darin bestehen, dass der Säugling noch sehr mangelhafte semantisch-interpretative Fähigkeiten für seinen sprachlichen Input aufweist. Dies könnte dazu führen, dass sich die Aufmerksamkeit beim Säugling stark auf formale Aspekte des sprachlichen Inputs fokussiert. Demgegenüber konzentriert sich der ältere Sprachlerner vielleicht von Beginn an stärker auf inhaltliche Aspekte und hat deshalb größere Probleme bei der Erfassung formaler Eigenschaften des sprachlichen Systems. Diese Überlegungen werden durch Untersuchungsbefunde zum Erwerb künstlicher Grammatiken bei Erwachsenen gestützt. Wie in Abschnitt 5.1.4 dargestellt, sind erwachsene Sprecher in Experimenten zum Erwerb künstlicher Grammatiken relativ erfolgreich in der Verwendung formaler Aspekte von Äußerungen einer Kunstsprache, um die Regeln der Kunstsprache zu lernen. Allerdings zeigen sich bei Erwachsenen Unterschiede in den Erwerbsstrategien von künstlichen Grammatiken in Abhängigkeit davon, ob den im Input verwendeten sprachlichen Einheiten Referenten zugewiesen worden waren oder nicht. War ein Referenzfeld vorhanden, orientierten sich die Lerner stärker an der semantischen Information als an formalen Aspekten des sprachlichen Inputs (Mori & Moeser, 1986). Dieses Muster deutet darauf hin, dass Unterschiede im Erwerbserfolg zwischen älteren Lernern und Kindern wahrscheinlich nicht darauf beruhen, dass Erwachsene über bestimmte Lernmechanismen, die der junge Erstspracherwerber nutzen kann, grundsätzlich nicht mehr verfügen. Vielmehr deutet sich beim älteren Lerner eine andere Gewichtung der Inputinformation zugunsten einer stärkeren Berücksichtigung semantischer Informationsstrategien bei der Verarbeitung sprachlicher Äußerungen an, eine Strategie, die - wie oben schon angedeutet - möglicherweise nicht so effizient für den Erwerb formaler Eigenschaften des sprachlichen Systems ist. Auch hier werden zukünftige Untersuchungen sicherlich mehr Klarheit bringen.

Referenzen

- Abercrombie, D. (1967) *Elements of general phonetics*. Edinburgh: University Press.
- Aitchinson, J. (1987) *Words in the mind*. Oxford: Blackwell.
- Aslin, R. N., Pisoni, D. B., Hennessy, B. L. & Perey, A. J. (1981) Discrimination of voice onset time by human infants: New findings and implications for the effects of early experience. *Child Development* 52, 1135-1145.
- Aslin, R. N., Woodward, J. Z., LaMendola, N. P. & Bever, T. G. (1996) Models of word segmentation in fluent maternal speech to infants. In: J. L. Morgan & K. Demuth (eds.) *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*. Mahwah: Erlbaum.
- Auer, P. & Uhmann, S. (1988) Silben- und akzentzählende Sprachen. *Zeitschrift für Sprachwissenschaft* 7, 214-259.
- Baayen, H., Piepenbrock, R. & van Rijn, H. (1993) *The CELEX Lexical Database*. Philadelphia, PA.: Linguistic Data Consortium.
- Bates, E. & Goodman, J. C. (1999) On the emergence of grammar from the lexicon. In B. MacWhinney (ed.) *The Emergence of Language*. Mahwah: Erlbaum.
- Bates, E. & MacWhinney, B. (1989) Functionalism and the Competition Model. In B. MacWhinney & E. Bates (eds.) *The Crosslinguistic Study of Sentence Processing*. Cambridge, MA: University Press.
- Bellugi, U., Wang, P. & Jernigan, T. (1994) Williams syndrome: An unusual neuropsychological profile. In S. Broman & J. Grafman (eds.) *Atypical cognitive deficits in developmental disorders. Implications for brain function*. Hillsdale: Erlbaum.
- Bernstein Ratner, N. (1986) Durational cues which mark clause boundaries in mother-child speech. *Journal of Phonetics* 14, 1303-1309.
- Bernstein Ratner, N. & Rooney, B. (2001) How accessible is the lexicon in motherese. In J. Weissenborn & B. Höhle (eds.) *Approaches to Bootstrapping: Phonological, Lexical, Syntactic and Neurophysiological Aspects of Early Language Acquisition. Vol. 1*. Amsterdam: Benjamins.
- Bertoncini, J. & Mehler, J. (1981) Syllables as units in infant speech perception. *Infant Behavior and Development* 4, 247-260.
- Bertoncini, J., Bijeljac-Babic, R., Blumstein, S. E. & Mehler, J. (1987) Discrimination in neonates of very short CV's. *Journal of the Acoustical Society of America* 82, 31-37.
- Bertoncini, J., Bijeljac-Babic, R., Jusczyk, P. W., Kennedy, L. J. & Mehler, J. (1988) An investigation of young infants' perceptual representations of speech sounds. *Journal of Experimental Psychology: General* 117, 21-33.
- Bertoncini, J., Floccia, C., Nazzi, T. & Mehler, J. (1995) Morae and syllables: Rhythmical basis of speech representations in neonates. *Language and Speech* 38, 311-329.
- Best, C. T. (1990) Adult perception of nonnative contrasts differing in assimilation to native phonological categories. *Journal of the Acoustical Society of America* 88, 177.
- Best, C. T. (1993) Emergence of language-specific constraints in perception of non-native speech: A window on early phonological development. In: B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. McNeilage & J. Morton (eds.) *Developmental Neurocognition: Speech and Face Processing in the First Year of Life*. Dordrecht: Kluwer.
- Best, C. T., McRoberts, G. W. & Sithole, N. M. (1988) Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance* 14, 345-360.

- Bialystok, E. & Hakuta, K. (1999) Confounded age: Linguistic and cognitive factors in age differences for second language acquisition. In D. Birdsong (ed.) *Second language acquisition and the critical period hypothesis*. Mahwah: Erlbaum.
- Bickerton, D. (1999) Creole languages, the language bioprogram hypothesis, and language acquisition. In W. C. Ritchie & T. K. Bhatia (eds.) *Handbook of Child Language Acquisition*. San Diego: Academic Press.
- Bijeljac-Babic, R., Bertoncini, J. & Mehler, J. (1993) How do four-day-old infants categorize multisyllabic utterances. *Developmental Psychology* 29, 711-721.
- Birdsong, D. & Molis, M. (2001) On the evidence for maturational constraints in second-language acquisition. *Journal of Memory and Language* 44, 235-249.
- Birdsong, D. (1999) Whys and why nots of the critical period hypothesis for second language acquisition. In D. Birdsong (ed.) *Second language acquisition and the critical period hypotheses*. Mahwah: Erlbaum.
- Bolinger, D. (1965) *Pitch accent and sentence rhythm. Forms of English: Accent, morpheme order*. Cambridge, MA: Harvard University Press.
- Borzone de Manrique, A. & Signorini, A. (1983) Segmental duration and rhythm in Spanish. *Journal of Phonetics* 11, 117-128.
- Bosch, L. & Sebastián-Gallés, N. (1997) Native-language recognition abilities in 4-month-old infants from monolingual and bilingual environments. *Cognition* 65, 33-69.
- Bosch, L. & Sebastián-Gallés, N. (2001) Evidence of early language discrimination in infants from bilingual environments. *Infancy* 2, 29-49.
- Bowerman, M. (1973) *Early syntactic development: A cross-linguistic study with special reference to Finnish*. Cambridge: University Press.
- Bowerman, M. & Levinson, S. C. (2001) (eds.) *Language acquisition and conceptual development*. Cambridge: University Press.
- Bradley, D., Sánchez-Casas, R. & García-Albea, J. (1993) The status of the syllable in the perception of Spanish and English. *Language and Cognitive Processes* 8, 197-233.
- Braine, M. D. S. (1992) What sort of innate structure is needed to 'bootstrap' into syntax? *Cognition* 45, 77-100.
- Brent, M. R. & Cartwright, T. A. (1996) Distributional regularity and phonotactic constraints are useful for segmentation. *Cognition* 61, 93-125.
- Brent, M. R. & Siskind, J. M. (2001) The role of exposure to isolated words in early vocabulary development. *Cognition* 81, B33-B44.
- Brown, R. (1973) *A first language: The early stages*. Cambridge, MA: Harvard University Press.
- Burnham, D. & Dodd, B. (1998) Familiarity and novelty preferences in infants' auditory-visual speech perception: Problems, factors, and a solution. In C. Rovee-Collier, L. P. Lipsitt & H. Hayne (eds.) *Advances in Infancy Research, Vol. 12*. Stamford: Alex Publishing Corporation. (S. 170-187)
- Cairns, P., Shillcock, R., Chater, N. & Levy, J. (1997) Bootstrapping word boundaries: A bottom-up corpus-based approach to speech segmentation. *Cognitive Psychology* 33, 111-153.
- Cassidy, K. W. & Kelly, M. H. (1991) Phonological information for grammatical category assignments. *Journal of Memory and Language* 30, 348-369.
- Chomsky, N. (1965) *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Chomsky, N. (1981) *Lectures on government and binding*. Dordrecht: Foris.
- Chomsky, N. (1986) *Knowledge of language. Its nature, origin and use*. New York: Praeger.

- Chomsky, N. (1999) On the nature, use, and acquisition of language. In W. C. Ritchie & T. K. Bhatia (eds.) *Handbook of Child Language Acquisition*. San Diego: Academic Press.
- Christiansen, M. H., Allen, J. & Seidenberg, M. S. (1998) Learning to segment speech using multiple cues: A connectionist model. *Language and Cognitive Processes* 12, 221-268.
- Christophe, A. & Morton, J. (1998) Is Dutch native English? Linguistic analysis by 2month-olds. *Developmental Science* 1, 215-219.
- Christophe, A., Dupoux, E., Bertoncini, J. & Mehler, J. (1994) Do infants perceive word boundaries? An empirical study of the bootstrapping of lexical acquisition. *Journal of the Acoustical Society of America* 95, 1570-1580.
- Christophe, A., Guasti, T., Nespor, M., Dupoux, E. & van Ooyen, B. (1997) Reflections on phonological bootstrapping: Its role for lexical and syntactic acquisition. *Language and Cognitive Processes* 12, 585-612.
- Church, K. (1987) Phonological parsing and lexical retrieval. *Cognition* 25, 53-69.
- Clahsen, H. (1982) *Spracherwerb in der Kindheit. Eine Untersuchung zur Entwicklung der Syntax bei Kleinkindern*. Tübingen: Narr.
- Clahsen, H. (1999) Linguistic perspectives on specific language impairment. In W. C. Ritchie & T. K. Bhatia (eds.) *Handbook of Child Language Acquisition*. San Diego: Academic Press.
- Clahsen, H. & Almazan-Hamilton, M. (1998) Syntax and morphology in Williams Syndrome. *Cognition* 68, 167-198.
- Clark, E. (1995) Later lexical development and word formation. In P. Fletcher & B. MacWhinney (eds.) *The Handbook of Child Language*. Cambridge: Blackwell.
- Cole, R. A. & Jakimik, J. (1980) A model of speech perception. In R. A. Cole (ed.) *Perception and production of fluent speech*. London: Academic Press.
- Colombo, J. & Bundy, R. S. (1983) Infant response to auditory familiarity and novelty. *Infant Behavior and Development* 6, 305-311.
- Cooper, W. E. & Paccia-Cooper, J. (1980) *Syntax and Speech*. Cambridge, MA: Harvard University Press.
- Cooper, W. E. & Sorenson, J. M. (1977) Fundamental frequency contours at syntactic boundaries. *Journal of the Acoustical Society of America* 62, 683-692.
- Cowan, N. (1991) Recurrent speech patterns as cues to the segmentation of multisyllabic sequences. *Acta Psychologica* 77, 121-135.
- Crain, S. (1991) Language acquisition in the absence of experience. *Behavioral and Brain Sciences* 14, 597-650.
- Crystal, T. H. & House, A. S. (1988) Segmental durations in connected-speech signals: Current results. *Journal of the Acoustical Society of America* 83, 1553-1573.
- Curtiss, S. (1977) *Genie: a psycholinguistic study of a modern-day "wild child"*. New York: Academic Press.
- Cutler, A. (1986) *Forbear* is a homophone: Lexical prosody does not constrain lexical access. *Language and Speech* 29, 201-220.
- Cutler, A. (1993) Phonological cues to open- and closed-class words in the processing of spoken sentences. *Journal of Psycholinguistic Research* 22, 109-131.
- Cutler, A. (1994) Segmentation problems, rhythmic solutions. *Lingua* 92, 81-104.
- Cutler, A. & Butterfield, S. (1992) Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language* 31, 218-236.
- Cutler, A. & Carter, D. M. (1987) The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language* 2, 133-142.

- Cutler, A. & Clifton, C. E. (1984) The use of prosodic information in word recognition. In: H. Bouma & D. G. Bouwhuis (eds.) *Attention and Performance X: Control of Language processes*. Hillsdale, N. J.: Erlbaum.
- Cutler, A. & Norris, D. (1988) The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance* 14, 113-121.
- Cutler, A., Dahan, D. & van Donselaar, W. (1997) Prosody in the comprehension of spoken language: A literature review. *Language and Speech* 2, 133-142.
- Cutler, A., Mehler, J., Norris, D. & Segui, J. (1986) The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language* 25, 385-400.
- Cutler, A., Mehler, J., Norris, D. & Segui, J. (1992) The monolingual nature of speech segmentation by bilinguals. *Cognitive Psychology* 24, 381-410.
- Dale, P. S. (1974) Hesitations in maternal speech. *Language and Speech* 17, 174-181.
- Dauer, R. M. (1983) Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics* 11, 51-62.
- De Boysson-Bardies, B. & Vihman, M. M. (1991) Adaptation to language: Evidence from babbling and first words in four languages. *Language* 67, 297-319.
- De Boysson-Bardies, B., Hallé, P., Sagart, L. & Durand, C. (1989) A cross-linguistic investigation of vowel formants in babbling. *Journal of Child Language* 16, 1-17.
- DeCasper, A. J. & Fifer, W. P. (1980) Of human bonding: Newborns prefer their mother's voice. *Science* 208, 1174-1176.
- Durieux, G. & Gillis, S. (2001) Predicting grammatical classes from phonological cues: An empirical test. In: J. Weissenborn & B. Höhle (eds.) *Approaches to Bootstrapping: Phonological, Lexical, Syntactic and Neurophysiological Aspects of Early Language Acquisition*. Amsterdam: Benjamins.
- Echols, C. H., Crowhurst, M. J. & Childers, J. B. (1997) The perception of rhythmic units in speech by infants and adults. *Journal of Memory and Language* 36, 202-225.
- Eilers, R. E. (1977) Context sensitive perception of naturally produced stop and fricative consonants by infants. *Journal of the Acoustical Society of America* 61, 1321-1336.
- Eilers, R. E., Wilson, E. R. & Moore, J. M. (1977) Developmental changes in speech discrimination in infants. *Journal of Speech and Hearing Research* 20, 766-780.
- Eimas, P. D. (1974) Auditory and linguistic processing of cues for place of articulation by infants. *Perception and Psychophysics* 16, 513-521.
- Eimas, P. D. (1975) Auditory and phonetic coding of the cues for speech. Discrimination of the [r-] distinction by young infants. *Perception and Psychophysics* 18, 341-347.
- Eimas, P. D. & Miller, J. L. (1980) Discrimination of the information for manner of articulation. *Infant Behavior and Development* 3, 367-375.
- Eimas, P. D., Siqueland, E. R., Jusczyk, P. & Vigorito, J. (1971) Speech perception in infants. *Science* 171, 303-306.
- Elman, J. L. (1993) Learning and development in neural networks: the importance of starting small. *Cognition* 48, 71-99.
- Elman, J. L. (1999) The emergence of language: A conspiracy theory. In B. MacWhinney (ed.) *The Emergence of Language*. Mahwah: Erlbaum.
- Elman, J., Bates, E., Johnson, M., Karmiloff-Smith, A., Parisi, D. & Plunkett, K. (1996) *Rethinking Innateness*. Cambridge, MA: The MIT Press
- Epstein, S. D., Flynn, S. & Martohardjono, G. (1996) Second language acquisition: Theoretical and experimental issues in contemporary research. *Behavioral and Brain Sciences* 19, 677-758.

- Fenson, L., Bates, E., Dale, P., Pethick, S., Reznick, S. & Thal, D. (1994) *Variability in Early Communicative Development*. Chicago: The University of Chicago Press.
- Fernald, A. (1985) Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development* 8, 181-195.
- Fernald, A. & Kuhl, P. (1987) Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development* 10, 279-293.
- Fernald, A. & McRoberts, G. W. (1996) Prosodic bootstrapping: A critical analysis of the argument and the evidence. In J. L. Morgan & K. Demuth (eds.) *Signal to syntax: Bootstrapping from Speech to Syntax in Early Language Acquisition*. Hillsdale, N.J.: Erlbaum.
- Fernald, A. & Simon, T. (1984) Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology* 20, 104-113.
- Fernald, A., Taeschner, T., Dunn, J., Papousek, M., De Boysson-Bardies, B. & Furui, I. (1989) A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language* 16, 477-501.
- Ferreira, F. (1993) Creation of prosody during sentence production. *Psychological Review* 100, 233-243.
- Ferreira, F. & Clifton, C. (1986) The independence of syntactic processing. *Journal of Memory and Language* 25; 348-368.
- Fifer, W. P. & Moon, C. (1989) Psychobiology of newborn auditory preferences. *Seminars in Perinatology* 13, 430-433.
- Finney, S. A., Protopapas, A. & Eimas, P. D. (1996) Attentional allocation to syllables in American English. *Journal of Memory and Language* 35, 893-909.
- Fisher, C. & Tokura, H. (1996) Prosody in speech to infants: Direct and indirect acoustic cues to syntactic structure. In: J. L. Morgan & K. Demuth (eds.) *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*. Mahwah: Erlbaum.
- Flege, J. E. (1989) Chinese subjects' perception of the word-final English /t/ - /d/ contrast: Before and after training. *Journal of the Acoustical Society of America* 86, 1684-1697.
- Fodor, J. A. & Bever, T. G. (1965) The psychological reality of linguistic segments. *Journal of Verbal Learning and Verbal Behavior* 4, 414-420.
- Fodor, J. A. & Garrett, M. F. (1967) Some syntactic determinants of sentential complexity. *Perception and Psychophysics*, 2, 289-296.
- Foss, D. J. & Blank, M. A. (1980) Identifying the speech codes. *Cognitive Psychology* 12, 1-31.
- Frauenfelder, U. H. & Peeters, G. (1990) Lexical segmentation in TRACE: An exercise in simulation. In: G. T. M. Altmann (ed.) *Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives*. Cambridge, MA: MIT-Press.
- Frazier, L. (1978) *On comprehending sentences: Syntactic parsing strategies*. Bloomington: Indiana University Linguistics Club.
- Frazier, L. & Clifton, C. (1996) *Construal*. Cambridge, MA: MIT Press.
- Frazier, L. & Fodor, J. D. (1978) The sausage machine: A new two-stage parsing model. *Cognition* 6, 291-325.
- Frazier, L. & Rayner, K. (1982) Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology* 14, 178-210.
- Friederici, A. D. (1983) Aphasic's perception of words in sentential context: Some real-time processing evidence. *Neuropsychologia* 21, 351-358.
- Friederici, A. D. (1985) Levels of processing and vocabulary types: Evidence from on-line comprehension in normals and agrammatics. *Cognition* 19, 133-166.

- Friederici, A. D. & Wessels, J. M. I. (1993) Phonotactic knowledge and its use in infant speech perception. *Perception & Psychophysics* 54, 287-295.
- Friederici, A. D., Steinhauer, K. & Pfeifer, E. (2002) Brain signatures of artificial language processing: Evidence challenging the critical period hypothesis. *PNAS* 99, 529-534.
- Garnica, O. (1977) Some prosodic and paralinguistic features of speech to young children. In: C. E. Snow & C. A. Ferguson (eds.) *Talking to children: Language input and acquisition*. Cambridge, MA: University Press.
- Garrett, M., Bever, T. G. & Fodor, J. A. (1965) The active use of grammar in speech perception. *Perception & Psychophysics* 1, 30-32.
- Gathercole, S. & Baddeley, A. D. (1993) *Working memory and language*. Hove: Erlbaum.
- Gerken, L.A. (1991) The metrical basis of children's subjectless sentences. *Journal of Memory and Language* 30, 431-451.
- Gerken, L.A. (1994a) A metrical template account of children's weak syllable omissions from multisyllabic words. *Journal of Child Language* 21, 565-584.
- Gerken, L.A. (1994b) Young children's representation of prosodic phonology: Evidence from English-speakers' weak syllable productions. *Journal of Memory and Language* 33, 19-38.
- Gerken, L.A. (1996) Phonological and distributional information in syntax acquisition. In: J. L. Morgan & K. Demuth (eds.) *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*. Mahwah: Erlbaum.
- Gerken, L.A. (1996) Prosodic structure in young children's language production. *Language* 72, 683-712.
- Gerken, L.A. (2001) Signal to Syntax: Building a Bridge. In: J. Weissenborn & B. Höhle (eds.) *Approaches to Bootstrapping: Phonological, Lexical, Syntactic and Neurophysiological Aspects of Early Language Acquisition. Vol. 1*. Amsterdam: Benjamins.
- Gerken, L.A. & McIntosh, B. (1993) Interplay of function morphemes and prosody in early language. *Developmental Psychology* 29, 447-457.
- Gerken, L.A., Jusczyk, P. W. & Mandel, D. R. (1994) When prosody fails to cue syntactic structure: Nine-month-olds' sensitivity to phonological vs. syntactic phrases. *Cognition* 51, 237-265.
- Gerken, L.A., Landau, B. & Remez, R. E. (1990) Function morphemes in young children's speech perception and production. *Developmental Psychology* 26, 204-216.
- Gleitman, L. R. (1990) The structural sources of verb meanings. *Language Acquisition* 1, 3-55.
- Gleitman, L. & Gleitman, H. (2001) Bootstrapping a first vocabulary. In J. Weissenborn & B. Höhle (eds.) *Approaches to Bootstrapping: Phonological, Lexical, Syntactic and Neurophysiological Aspects of Early Language Acquisition. Vol. 1*. Amsterdam: Benjamins.
- Gleitman, L. R. & Wanner, E. (1982). Language acquisition: the state of the state of the art. In: E. Wanner & L. R. Gleitman (eds.) *Language Acquisition: The State of the Art*. Cambridge, MA: Cambridge University Press.
- Gold, E. M. (1967) Language identification in the limit. *Information and Control* 16, 447-474.
- Goldinger, S. D., Luce, P. A. & Pisoni, D. B. (1989) Priming lexical neighbors of spoken words: Effects of competition and inhibition. *Journal of Memory and Language* 28, 501-518.
- Goldin-Meadow, S. & Mylander, C. (1998) Spontaneous sign systems created by deaf children in two cultures. *Nature* 391, 279-281.
- Goldin-Meadow, S., Mylander, C. & Butcher, C. (1995) The resilience of combinatorial structure at the word level: Morphology in self-styled gesture systems. *Cognition* 56, 195-262.
- Goldman-Eisler, F. (1972) Pauses, clauses and sentences. *Language and Speech* 15, 103-113.
- Goldsmith, J. (1990) *Autosegmental and Metrical Phonology*. Oxford: Basil Blackwell.

- Golinkoff, R. M., Hirsh-Pasek, K. & Schweisguth, M. A. (2001) A reappraisal of young children's knowledge of grammatical morphemes. In: J. Weissenborn & B. Höhle (eds.) *Approaches to Bootstrapping: Phonological, Lexical, Syntactic and Neurophysiological Aspects of Early Language Acquisition. Vol. 1.* Amsterdam: Benjamins.
- Golinkoff, R. M., Hirsh-Pasek, K., Cauley, K. M. & Gordon, L. (1987) The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language* 14, 23-45.
- Gomez, R. L. & Gerken, L. A. (1999) 11-month-olds are sensitive to structure in an artificial grammar. *Cognition* 70, 109-135.
- Goodluck, H. (1991) *Language Acquisition: A Linguistic Introduction.* Oxford: Blackwell.
- Goodsitt, J., Morgan, J. L. & Kuhl, P. K. (1993) Perceptual strategies in prelingual speech segmentation. *Journal of Child Language* 20, 229-252.
- Gordon, P. (1985) Evaluating the semantic categories hypothesis: The case of the count/mass distinction. *Cognition* 20, 209-242.
- Green, T. R. G. (1979) The necessity of syntax markers: Two experiments with artificial languages. *Journal of Verbal Learning and Verbal Behavior* 18, 481-496.
- Grieser, D. & Kuhl, P. (1988) Maternal speech to infants in a tonal language: Support for universal prosodic features in motherese. *Developmental Psychology* 24, 14-20.
- Grosjean, F. & Collins, M. (1979) Breathing, pausing and reading. *Phonetics* 36, 98-114.
- Grosjean, F., Grosjean, L. & Lane, H. (1979) The patterns of silence: Performance structures in sentence production. *Cognitive Psychology* 11, 58-81.
- Guasti, M. T., Christophe, A., van Ooyen, B. & Nespors, M. (2001) Pre-lexical setting of the head complement parameter. In: J. Weissenborn & B. Höhle (eds.) *Approaches to Bootstrapping: Phonological, Lexical, Syntactic and Neurophysiological Aspects of Early Language Acquisition. Vol. 1.* Amsterdam: Benjamins.
- Hahne, A. (2001) What's different in second-language processing? Evidence from event-related brain potentials. *Journal of Psycholinguistic Research* 30, 251-266.
- Haith, M. M. & Benson, J. B. (1998) Infant cognition. In W. Damon (ed.) *Handbook of Child Psychology. Vol. 2: Cognition, Perception, and Language.* New York: Wiley.
- Hakes, D. T. (1972) Effects of reducing complement constructions on sentence comprehension. *Journal of Verbal Learning and Verbal Behavior* 11, 278-286.
- Hakes, D. T. & Cairns, H. S. (1970) Sentence comprehension and relativ pronouns. *Perception and Psychophysics* 8, 5-8.
- Hakes, D. T., Evans, J. S. & Brannon, L. L. (1976) Understanding sentences with relativ clauses. *Memory and Cognition* 4, 283-290.
- Harrington, J., Watson, G., & Cooper, M. (1989) Word boundary detection in broad class and phoneme strings. *Computer Speech and Language* 3, 367-382.
- Harris, Z. S. (1955) From phoneme to morpheme. *Language* 31, 190-222.
- Hauser, M. D., Newport, E. L. & Aslin, R. N. (2001) Segmentation of the speech stream in a non-human primate: Statistical learning in cotton-top tamarins. *Cognition* 78; B53-B64.
- Hawkins, P. R. (1971) The syntactic location of hesitation pauses. *Language and Speech* 14, 277-288.
- Hayes, J. R. & Clark, H. H. (1970) Experiments on the segmentation of an artificial speech analogue. In J. R. Hayes (ed.) *Cognition and the development of language.* New York: Wiley.
- Hennon, E., Hirsh-Pasek, K. & Golinkoff, R. M. (2000) Die besondere Reise vom Fötus zum sprachwerbenden Kind. In H. Grimm (ed.) *Enzyklopädie der Psychologie.* Göttingen: Hogrefe.
- Hillenbrand, J. M., Minifie, F. D. & Edwards, T. J. (1979) Tempo of spectrum change as a cue in speech sound discrimination by infants. *Journal of Speech and Hearing Research* 22, 147-165.

- Hirsh-Pasek, K. & Golinkoff, R. M. (1996) *The origins of grammar: Evidence from early language comprehension*. Cambridge, MA: MIT Press.
- Hirsh-Pasek, K., Kemler Nelson, D. G., Jusczyk, P. W., Wright Cassidy, K., Druss, B. & Kennedy, L. (1987) Clauses are perceptual units for young infants. *Cognition* 26, 269-286.
- Höhle, B. & Schriefers, H. (1995) Ambisyllabizität im Deutschen: Psycholinguistische Evidenz. In: P. Baerentzen (Hrsg.) *Aspekte der Sprachbeschreibung*. Tübingen: Niemeyer.
- Höhle, B. & Weissenborn, J. (1998) Sensitivity to closed-class-elements in preverbal children. In In A. Greenhill, M. Hughes, H Littlefield, & H. Walsh (eds.) *Proceedings of the 22nd Annual Boston Conference on Language Development*, Somerville, Cascadilla Press.
- Höhle, B. & Weissenborn, J. (1999) Discovering grammar: Prosodic and morpho-syntactic aspects of rule formation in First Language Acquisition. In: A. D. Friederici & R. Menzel (eds.) *Learning: Rule Extraction and Representation*. Berlin: de Gruyter.
- Höhle, B. & Weissenborn, J. (2000) Lauter Laute? Lautsegmente und Silben in der Sprachperzeption und im Spracherwerb. In R. Thieroff, M. Tamrat, N. Furhop & O. Teuber, O. (Hrsg.) *Deutsche Grammatik in Theorie und Praxis*. Tübingen: Niemeyer.
- Höhle, B. & Weissenborn, J. (in press) German-learning infants' ability to detect unstressed closed-class elements in continuous speech. *Developmental Science*.
- Höhle, B., Hofmann, M. & Friederici, A. D. (1995) Verarbeitung prosodischer Information in der Aphasie. Poster auf der 22. Jahrestagung der Arbeitsgemeinschaft für Aphasieforschung und –behandlung. Magdeburg.
- Höhle, B., Weissenborn, J., Schmitz, M. & Ischebeck, A. (2001) Word order regularities: The role of prosodic information for early parameter setting. In J. Weissenborn & B. Höhle (eds.) *Approaches to Bootstrapping: Phonological, Lexical, Syntactic and Neurophysiological Aspects of Early Language Acquisition, Vol. 1*. Amsterdam: Benjamins.
- Hohne, E. A. & Jusczyk, P. W. (1994) Two-month-old infants' sensitivity to allophonic differences. *Perception and Psychophysics* 56, 613-623.
- Hollich, G., Hirsh-Pasek, K. & Golinkoff, R. M. (1998) Introducing the 3D intermodal preferential looking paradigm: A new method to answer an age-old question. In C. Rovee-Collier, L. P. Lipsitt & H. Hayne (eds.) *Advances in Infancy Research, Vol. 12*. Stamford: Alex Publishing Corporation.
- Horowitz, F. D. (1975) Infant attention and discrimination: Methodological and substantive issues. *Monographs of the Society for Research in Child Development* 39.
- Houston, D. M. & Jusczyk, P. W. (2000) The role of talker-specific information in word segmentation by infants. *Journal of Experimental Psychology: Human Perception and Performance* 26, 1570-1582.
- Houston, D. M. & Jusczyk, P. W. (2001) Infants' long-term memory for words and voices. Ms. Indiana University Medical School.
- Houston, D. M., Jusczyk, P. W., Kuijpers, C., Coolen, R. & Cutler, A. (2000) Crosslanguage word segmentation by 9-month-olds. *Psychonomic Bulletin & Review* 7, 504-509.
- Hunter, M. A. & Ames, E. W. (1988) A multifactor model of infant preferences for novel and familiar stimuli. In L. P. Lipsitt & C. Rovee-Collier (eds.) *Advances in Infancy Research, Vol. 5*. Stamford: Alex Publishing Corporation.
- Hunter, M. A., Ames, E. W. & Koopman, R. (1983) Effects of stimulus complexity and familization time on infant preferences for novel and familiar stimuli. *Developmental Psychology* 19, 338-352.
- Jessen, M., Marasek, K., Schneider, K. & Clahßen, K. (1995) Acoustic correlates of word stress and the tense/lax opposition in the vowel system of German. *Proceedings of the XIIIth International Congress of Phonetic Sciences*. Vol. 4. Stockholm. Schweden.
- Johnson, E. K. & Jusczyk, P. W. (2001) Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory and Language* 44, 548-567.

- Johnson, J. S. & Newport, E. L. (1989) Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology* 21, 60-99.
- Jusczyk, P. W. (1977) Perception of syllable-final stops by two-month-old infants. *Perception & Psychophysics* 21, 450-454.
- Jusczyk, P. W. (1997) *The Discovery of Spoken Language*. Cambridge, MA: MIT-Press.
- Jusczyk, P. W. & Aslin, R. N. (1995) Infants' detection of sound patterns of words in fluent speech. *Cognitive Psychology* 29, 1-23.
- Jusczyk, P. W. & Hohne, E. A. (1997) Infant's memory for spoken words. *Science* 277, 1984-1986.
- Jusczyk, P. W. & Krumhansl, C. L. (1993) Pitch and rhythmic patterns affecting infants' sensitivity to musical phrase structure. *Journal of Experimental Psychology: Human Perception and Performance* 19, 627-640.
- Jusczyk, P. W. & Thompson, E. (1978) Perception of a phonetic contrast in multisyllabic utterances by 2-month-old infants. *Perception & Psychophysics* 23, 105-109.
- Jusczyk, P. W., Copan, H. & Thompson, E. (1978) Perception by two-month-olds of glide contrasts in multisyllabic utterances. *Perception and Psychophysics* 24, 515-520.
- Jusczyk, P. W., Cutler, A. & Redanz, N. (1993) Preference for the predominant stress patterns of English words. *Child Development* 64, 675-687.
- Jusczyk, P. W., Friederici, A. D., Wessels, J., Svenkerud, V. Y. & Jusczyk, A. M. (1993) Infants' sensitivity to the sound patterns of native language words. *Journal of Memory and Language* 32, 402-420.
- Jusczyk, P.W., Hirsh-Pasek, K., Kemler Nelson, D.G., Kennedy, L.J., Woodward, A., & Piwoz, J. (1992) Perception of acoustic correlates of major phrasal units by young infants. *Cognitive Psychology* 24, 252-293.
- Jusczyk, P. W., Hohne, E. A. & Bauman, A. (1999) Infants' sensitivity to allophonic cues for word segmentation. *Perception & Psychophysics* 61, 1465-1476.
- Jusczyk, P. W., Houston, D. M. & Newsome, M. (1999) The beginnings of word segmentation in English-learning infants. *Cognitive Psychology* 39, 159-207.
- Jusczyk, P. W., Luce, P. A. & Charles-Luce, J. (1994) Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language* 33, 630-645.
- Jusczyk, P. W., Pisoni, D. B. & Mullenix, J. (1992) Some consequences of stimulus variability on speech processing by 2-month old infants. *Cognition* 43, 253-291.
- Karmiloff-Smith, A. (1992) *Beyond Modularity*. Cambridge, MA: MIT Press.
- Karmiloff-Smith, A., Tyler, L. K., Voice, K., Sims, K., Udwin, O., Howlins, P. & Davies, M. (1998) Linguistic dissociations in Williams syndrome: evaluating receptive syntax in on-line and off-line tasks. *Neuropsychologia* 36, 343-351.
- Katz, N., Baker, E. & Macnamara, J. (1974) What's in a name? A study of how children learn common and proper names. *Child Development* 45, 469-473.
- Kelly, M. (1996) The role of phonology in grammatical category assignments. In: J. L. Morgan & K. Demuth (eds.) *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Language Acquisition*. Mahwah: Erlbaum.
- Kelly, M. H. & Bock, J. K. (1988) Stress in time. *Journal of Experimental Psychology: Human Perception and Performance* 14, 389-403.
- Kelly, M. H. (1992) Using sound to solve syntactic problems: The role of phonology in grammatical category assignments. *Psychological Review* 99, 349-364.
- Kemler Nelson, D. G., Hirsh-Pasek, K., Jusczyk, P. W. & Wright-Cassidy, K. (1989) How prosodic cues in motherese might assist language learning. *Journal of Child Language* 16, 55-68.

- Klatt, D. H. (1975) Vowel lengthening is syntactically determined in a connected discourse. *Journal of Phonetics* 3, 129-140.
- Klatt, D. H. (1976) Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America* 59, 1203-1221.
- Kluender, K. R., Lotto, A. J., Holt, L. L. & Bloedel, S. L. (1998) Role of experience for language-specific functional mappings of vowel sounds. *Journal of the Acoustical Society of America* 104, 3568-3582.
- Köpcke, K.-M. (1982) *Untersuchungen zum Genusssystem der deutschen Gegenwartssprache*. Tübingen: Niemeyer.
- Köpcke, K.-M. & Zubin, D. (1984) Sechs Prinzipien für die Genuszuweisung im Deutschen: Ein Beitrag zur natürlichen Klassifikation. *Linguistische Berichte* 92, 26-50.
- Krumhansl, C. L. & Jusczyk, P. W. (1990) Infants' perception of phrase structure in music. *Psychological Science* 1, 70-73.
- Kuhl, P. K. (1979) Speech perception in early infancy: Perceptual constancy for spectrally dissimilar vowel categories. *Journal of the Acoustical Society of America* 66, 1668-1679.
- Kuhl, P. K. (1983) Perception of auditory equivalence classes for speech in early infancy. *Infant Behavior and Development* 6, 263-285.
- Kuhl, P. K. (1993) Innate predispositions and the effects of experience in speech perception: The native language magnet theory. In: B. de Boysson-Bardies, S. de Schonen, P. W. Jusczyk, P. McNeilage & J. Morton (eds.) *Developmental Neurocognition: Speech and Face Processing in the First Year of Life*. Dordrecht: Kluwer.
- Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N. & Lindblom, B. (1992) Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* 255, 606-608.
- Lasky, R. E., Syrda-Lasky, A. & Klein, R. E. (1975) VOT discrimination by four to six and a half month old infants from Spanish environments. *Journal of Experimental Child Psychology* 20, 215-225.
- Lecanuet, J.-P. & Granier-Deferre, C. (1993) Speech stimuli in the fetal environment. In: B. de Boysson-Bardies, S. de Schonen, P. W. Jusczyk, P. McNeilage & J. Morton (eds.) *Developmental Neurocognition: Speech and Face Processing in the First Year of Life*. Dordrecht: Kluwer.
- Lehiste, I. (1973) Phonetic disambiguation of syntactic ambiguity. *Glossa* 7, 107-222.
- Lehiste, I., Olive, J. P. & Streeter, L. A. (1976) Role of duration in disambiguating syntactically ambiguous sentences. *Journal of the Acoustical Society of America* 60, 1199-1202.
- Lenneberg, E. (1967) *Biological Foundations of Language*. New York: Wiley.
- Levitt, A., Jusczyk, P. W., Murray, J. & Carden, G. (1988) The perception of place of articulation contrasts in voiced and voiceless fricatives by two-month-old infants. *Journal of Experimental Psychology: Human Perception and Performance* 14, 361-368.
- Levy, Y. (1983) It's frogs all the way down. *Cognition* 15, 75-93.
- Lieberman, A. M., Cooper, P. C., Shankweiler, D. P. & Studdert-Kennedy, M. G. (1967) Perception of the speech code. *Psychological Review* 74, 431-461.
- Lieberman, A. M., Harris, K. H., Kinney, J. A. & Lane, H. L. (1961) The discrimination of relative-onset time of the components of certain speech and non-speech patterns. *Journal of Experimental Psychology* 61, 379-388.
- Lieberman, I. Y., Shankweiler, D., Fischer, W. & Carter, B. (1974) Explicit syllable and phoneme segmentation in the young child. *Journal of Experimental Child Psychology* 18, 201-212.
- Lieberman, P., Katz, W., Jongman, A., Zimmerman, R. & Miller, M. (1985) Measures of the sentence intonation of read and spontaneous in American English. *Journal of the Acoustical Society of America* 77, 649-657.

- Lively, S. E., Pisoni, D. B. & Goldinger, S. D. (1994) Spoken word recognition: Research and theory. In: M. A. Gernsbacher (ed.) *Handbook of Psycholinguistics*. San Diego: Academic Press.
- Lleó, C., & Demuth, K. (1999) Prosodic constraints on the emergence of grammatical morphemes: Cross-linguistic evidence from Germanic and Romance languages In *Proceedings of the 23rd Annual Boston Conference on Language Development*, Somerville: Cascadilla Press.
- Logan, J. S., Lively, S. E. & Pisoni, D. B. (1989) Training Japanese listeners to identify /r/ and /l/. *Journal of the Acoustical Society of America* 85, 137-138.
- Lust, B. (1999) Universal grammar: The strong continuity hypothesis in first language acquisition. In W. C. Ritchie & T. K. Bhatia (eds.) *Handbook of Child Language Acquisition*. San Diego: Academic Press.
- MacLay, H. & Osgood, C. E. (1959) Hesitation phenomena in spontaneous English speech. *Word* 15, 19-44.
- MacWhinney, B., Leinbach, J., Taraban, R. & McDonald, J. (1989) Language learning: Cues or rules? *Journal of Memory and Language* 28, 255-277.
- Mandel, D. R., Jusczyk, P. W. & Kemler Nelson, D. G. (1994) Does sentential prosody help infants to organize and remember speech information? *Cognition* 53, 155-180.
- Mandel, D. R., Jusczyk, P. W. & Pisoni, D. B. (1995) Infants' recognition of the sound patterns of their own names. *Psychological Science* 6, 315-318.
- Mandel, D. R., Kemler Nelson, D. G. & Jusczyk, P. W. (1996) Infants remember the order of words in a spoken sentence. *Cognitive Development* 11, 181-196.
- Marantz, A. P. (1984) *On the nature of grammatical relations*. Cambridge, MA: MIT Press.
- Maratsos, M. (1982) The child's construction of grammatical categories. In E. Wanner & L. R. Gleitman (eds.) *Language acquisition: the state of the art*. Cambridge: University Press.
- Maratsos, M. (1988) The acquisition of formal word classes. In Y. Levy, I. M. Schlesinger & M. D. S. Braine (eds.) *Categories and processes in language acquisition*. Hillsdale: Erlbaum
- Maratsos, M. & Chalkley, M. A. (1980) The internal language of children's syntax: The ontogenesis and representation of syntactic categories. In K. Nelson (ed.) *Children's language*, Vol. 2. New York: Gardner Press.
- Marcus, G. F., Vijayan, S., Bandi Rao, S. & Vishton, P. M. (1999) Rule learning by seven-month-old infants. *Science* 283, 77-80.
- Markman, E. M. (1994) Constraints on word meaning in early language acquisition. In L. Gleitman & B. Landau (eds.) *Lexical acquisition*. Amsterdam: Elsevier.
- Marler, P. (1991) The instinct to learn. In S. Carey & R. Gelman (eds.) *The epigenesis of mind: Essays on Biology and Cognition*. Hillsdale: Erlbaum.
- Marler, P. & Peter, S. (1987) A sensitive period for song acquisition in the song sparrow, *Melospiza melodia*: a case of age limited learning. *Ethology* 76, 89-100.
- Marslen-Wilson, W. D. (1987) Functional parallelism in spoken word-recognition. *Cognition* 25, 71-102.
- Marslen-Wilson, W. D. (1990) Activation, competition, and frequency in lexical access. In: G. T. M. Altmann (ed.) *Cognitive models of speech processing: Psycholinguistic and computational perspectives*. Cambridge, MA: MIT Press.
- Marslen-Wilson, W. S., Tyler, L. K., Warren, P., Grenier, P. & Lee, C. S. (1992) Prosodic effects in minimal attachment. *The Quarterly Journal of Experimental Psychology* 45, 73-87.
- Mattys, S. L. & Jusczyk, P. W. (2001a) Phonotactic cues for segmentation of fluent speech by infants. *Cognition* 78, 91-121.
- Mattys, S. L. & Jusczyk, P. W. (2001b) Do infants segment words of recurring contiguous patterns? *Journal of Experimental Psychology: Human Perception and Performance* 27, 644-655.

- Mattys, S. L., Jusczyk, P. W., Luce, P. A. & Morgan, J. L. (1999) Phonotactic and prosodic effects on word segmentation in infants. *Cognitive Psychology* 38, 465-494.
- Maye, J., Werker, J. F. & Gerken, L.A. (2002) Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* 82, B101-B111.
- Mazuka, R. (1996) Can a grammatical parameter be set before the first word? Prosodic contributions to early setting of a grammatical parameter. In J. L. Morgan & K. Demuth (eds.) *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*. Mahwah: Erlbaum.
- Mazuka, R. & Lust, B. (1990) On parameter setting and parsing: Predictions for cross-linguistic differences in adult and processing. In L. Frazier & J. de Villiers (eds.) *Language processing and language acquisition*. Dordrecht: Kluwer.
- McClelland, J. & Elman, J. (1986) The TRACE model of speech perception. *Cognitive Psychology* 18, 1-86.
- McDaniel, D., McKee, C. & Smith Cairns, H. (1996) (eds.) *Methods for Assessing Children's Syntax*. Cambridge, MA: MIT Press.
- McKee, C. & Iwasaki, N. (2001) Lemma structure in language learning: Comments on representation and realization. In J. Weissenborn & B. Höhle (eds.) *Approaches to Bootstrapping: Phonological, Lexical, Syntactic and Neurophysiological Aspects of Early Language Acquisition. Vol. 1*. Amsterdam: Benjamins.
- McQueen, J. M. (1998) Segmentation of continuous speech using phonotactics. *Journal of Memory and Language* 39, 21-46.
- McQueen, J. M., Norris, D. & Cutler, A. (1994) Competition in spoken word recognition: Spotting words in other words. *Journal of Experimental Psychology: Learning, Memory and Cognition* 20, 621-638.
- McQueen, J. M., Otake, T. & Cutler, A. (2001) Rhythmic cues and possible-word constraints in Japanese speech segmentation. *Journal of Memory and Language* 45, 103-132.
- Mecklinger, A., Schriefers, H., Steinhauer, K. & Friederici, A. D. (1995) Processing relative clauses varying on syntactic and semantic dimensions: An analysis with event-related potentials. *Memory and Cognition* 23, 477-494.
- Mehler, J., Dommergues, J. Y., Frauenfelder, U. & Segui, J. (1981) The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior* 20, 298-305.
- Mehler, J., Dupoux, E., Nazzi, T. & Dehaene-Lambertz, G. (1996) Coping with linguistic diversity: The infant's viewpoint. In J. L. Morgan & K. D. Demuth (eds.) *Signal to syntax: Bootstrapping from speech to syntax in early language acquisition*. Hillsdale, N.J.: Erlbaum.
- Mehler, J., Jusczyk, P. W., Lambertz, G., Halsted, G., Bertoncini, J. & Amiel-Tison, C. (1988) A precursor of language acquisition in young infants. *Cognition* 29, 143-178.
- Meier R. P. & Bower, G. H. (1986) Semantic reference and phrasal grouping in the acquisition of a miniature phrase structure language. *Journal of Memory and Language* 25, 492-505.
- Meints, K., Plunkett, K. & Harris, P. L. (1999) When does an ostrich become a bird? The role of typicality in early word comprehension. *Developmental Psychology* 35, 1072-1078.
- Meisenburg, T. & Selig, M. (1998) *Phonetik und Phonologie des Französischen*. Stuttgart: Klett.
- Miceli, G., Mazzuchi, A., Menn, L. & Goodglass, H. (1983) Contrasting cases of Italian agrammatic aphasia without comprehension disorder. *Brain and Language* 19, 65-97.
- Miller, J. L. & Eimas, P. D. (1983) Studies on the categorization of speech by infants. *Cognition* 13, 135-165.
- Mintz, T. (1996) *The roles of linguistic input and innate mechanisms in children's acquisition of grammatical categories*. PH.D. dissertation. University of Rochester, Rochester, NY.

- Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A. M., Jenkins, J. J. & Fujimura, O. (1975) An effect of linguistic experience: The discrimination of /r/ and /l/ by native speakers of Japanese and English. *Perception and Psychophysics* 18, 331-340.
- Moffitt, A. R. (1971) Consonant cue perception by twenty-to-twenty-four-week old infants. *Child Development* 42, 717-731.
- Moon, C., Panneton-Cooper, R. & Fifer, W. P. (1993) Two-day-olds prefer their native language. *Infant Behavior and Development* 16, 495-500.
- Morais, J., Bertelson, P., Cary, L. & Alegria, J. (1986) Literacy training and speech segmentation. *Cognition* 24, 45-64.
- Morais, J., Cary, L., Alegria, J. & Bertelson, P. (1979) Does awareness of speech as a sequence of phones arise spontaneously? *Cognition* 7, 323-331.
- Morgan, J. L. (1986) *From simple input to complex grammar*. Cambridge, MA: MIT Press.
- Morgan, J. L. (1994) Converging measures of speech segmentation in preverbal infants. *Infant Behavior and Development* 17, 389-403.
- Morgan, J. L. (1996) A rhythmic bias in preverbal speech segmentation. *Journal of Memory and Language* 35, 666-688.
- Morgan, J. L. & Saffran, J. R. (1995) Emerging integration of sequential and suprasegmental information in preverbal speech segmentation. *Child Development* 66, 911-936.
- Morgan, J. L., Meier, R. P. & Newport, E. L. (1987) Structural packaging in the input to language learning: Contributions of prosodic and morphological marking of phrases to the acquisition of language. *Cognitive Psychology* 19, 498-550.
- Morgan, J. L., Shi, R. & Allopenna, P. (1996) Perceptual bases of rudimentary grammatical categories: Toward a broader conceptualization of bootstrapping. In: J. L. Morgan & K. Demuth (eds.) *From Signal to Syntax: Bootstrapping from Speech to Grammar in Early Language Acquisition*. Mahwah: Erlbaum.
- Mori, K. & Moeser, S. D. (1983) The role of syntax markers and semantic referents in learning an artificial language. *Journal of Verbal Learning and Verbal Behavior* 22, 701-718.
- Morse, P. A. (1972) The discrimination of speech and nonspeech stimuli in early infancy. *Journal of Experimental Child Psychology* 13, 477-492.
- Myers, J., Jusczyk, P. W., Kemler Nelson, D. G., Charles Luce, J., Woodward, A. & Hirsh-Pasek, K. (1996) Infants' sensitivity to word boundaries in fluent speech. *Journal of Child Language* 23, 1-30.
- Naigles, L. (1990) Children use syntax to learn verb meaning. *Journal of Child Language* 17, 357-374.
- Nakatani, L. H. & Dukes, K. D. (1977) Locus of segmental cues for word juncture. *Journal of the Acoustical Society of America* 62, 714-719.
- Nazzi, T., Bertoncini, J. & Mehler, J. (1998) Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance* 24, 756-766.
- Nazzi, T., Jusczyk, P. W. & Johnson, E. K. (2000) Language discrimination by English-learning 5-month-olds: Effects of rhythm and familiarity. *Journal of Memory and Language* 43, 1-19.
- Nazzi, T., Kemler Nelson, D. G., Jusczyk, P. W. & Jusczyk, A. M. (2000) Six-month-olds' detection of clauses embedded in continuous speech: Effects of prosodic well-formedness. *Infancy* 1, 123-147.
- Nespor, M. & Vogel, I. (1986) *Prosodic Phonology*. Dordrecht: Foris.
- Nespor, M., Guasti M. T. & Christophe, A. (1996) Selecting word order: the Rhythmic Activation Principle. In: U. Kleinhenz (ed.) *Interfaces in Phonology*. Berlin: Akademie Verlag.

- Newport, E. L. (1977) Motherese: The speech of mothers to young children. In: N. Castelan, D. P. Pisoni & G. Potts (eds.) *Cognitive Theory, Vol. 2*. Hillsdale, NJ: Erlbaum.
- Newport, E. L. (1988) Constraints on learning and their role in language acquisition: Studies of the acquisition of American sign language. *Language Sciences* 10, 147-172.
- Newport, E. L. (1990) Maturational constraints on language learning. *Cognitive Science* 14, 11-28.
- Newport, E. L. (1991) Contrasting concepts of the critical period for language. In S. Carey & R. Gelman (eds.) *The epigenesis of mind: Essays on Biology and Cognition*. Hillsdale: Erlbaum.
- Newport, E. L. & Aslin, R. N. (2000) Innately constrained learning: Blending old and new approaches to language acquisition. In C. Howell, S. A. Fish & T. Keith-Lucas (eds.) *Proceedings of the 24th Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press.
- Norris, D. (1994) Shortlist: A connectionist model of continuous speech recognition. *Cognition* 52, 189-234.
- Norris, D., McQueen, J. M. & Cutler, A. (1995) Competition and segmentation in spoken word recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition* 21, 1209-1228.
- Norris, D., McQueen, J. M., Cutler, A. & Butterfield, S. (1997) The possible-word constraint in the segmentation of continuous speech. *Cognitive Psychology* 34, 191-243.
- O'Grady, W. (1999) The acquisition of syntactic representations: A general nativist approach. In W. C. Ritchie & T. K. Bhatia (eds.) *Handbook of Child Language Acquisition*. San Diego: Academic Press.
- Onishi, K. H., Chambers, K. E. & Fisher, C. (2002) Learning phonotactic constraints from brief auditory experience. *Cognition* 83, B13-B23.
- Osterhout, L., Holcomb, P. J. & Swinney, D. A. (1994) Brain potentials elicited by garden-path sentences: Evidence of the application of verb information during parsing. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 4, 786-803.
- Otake, T., Hatano, G., Cutler, A. & Mehler, J. (1993) Mora or syllable? Speech segmentation in Japanese. *Journal of Memory and Language* 32, 258-278.
- Pallier, C., Sebastian-Gallés, N., Felguera, T., Christophe, A. & Mehler, J. (1993) Attentional allocation within the syllabic structure of spoken words. *Journal of Memory and Language* 32, 373-389.
- Pegg, J. W., Werker, J. F. & McLeod, P. J. (1992) Preference for infant-directed over adult-directed speech: Evidence from 7-week-old infants. *Infant Behavior and Development* 15, 325-345.
- Penner, Z. & Weisenborn, J. (1996) Strong continuity, parameter setting and the trigger hierarchy. On the acquisition of the DP in Bernese Swiss German and High German. In: H. Clahsen (ed.) *Generative Perspectives on Language Acquisition: Empirical Findings, Theoretical Considerations, Crosslinguistic Comparisons*. Amsterdam: Benjamins.
- Peters, A. (1985) Language segmentation: Operating principles for the perception and analysis of language. In D. I. Slobin (ed.) *The Crosslinguistic Study of Language Acquisition. Vol. 2: Theoretical Issues*. Hillsdale, NJ: Erlbaum.
- Phillips, J. (1973) Syntax and vocabulary of mothers' speech to young children: age and sex comparisons. *Child Development* 44, 182-185.
- Piaget J. (1980) The psychogenesis of knowledge and its epistemological significance. In M. Piatelli-Palmarini (ed.) *Language and Learning*. London: Routledge & Kegan Paul.
- Pinker, S. (1984). *Language Learnability and Language Development*. Cambridge, MA: University Press.
- Pinker, S. (1987) The bootstrapping problem in language acquisition. In B. MacWhinney (ed.) *Mechanisms of Language Acquisition*. Hillsdale: Erlbaum.
- Pisoni, D. B. (1973) Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Perception and Psychophysics* 13, 253-260.

- Polka, L. & Werker, J. F. (1994) Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance* 20, 421-435.
- Price, P., Ostendorf, M., Shattuck-Hufnagel, S. & Fong, C. (1991) The use of prosody in syntactic disambiguation. *Journal of the Acoustical Society of America* 90, 2956-2970.
- Ramus, F., Nespors, M. & Mehler, J. (1999) Correlates of linguistic rhythm in the speech signal. *Cognition* 73, 265-292.
- Ratner, N. & Pye, C. (1984) Higher pitch in BT is not universal: acoustic evidence from Quiche Mayan. *Journal of Child Language* 2, 515-522.
- Read, C. Zhang, Y-F., Nie, H.-Y. & Ding, B.-Q. (1986) The ability to manipulate speech sounds depends on knowing alphabetic writing. *Cognition* 24, 31-44.
- Roach, P. (1982) On the distinction between 'stress-timed' and 'syllable-timed' languages. In: D. Crystal (ed.) *Linguistic controversies*. London: Arnold.
- Roeper, T. & Weissenborn, J. (1990) How to make parameters work. In L. Frazier & J. de Villiers (eds.) *Language processing and language acquisition*. Dordrecht: Kluwer.
- Rose, S. A., Gottfried, A. W., Melloy-Carminar, P. & Bridger, W. H. (1982) Familiarity and novelty preferences in infant recognition memory: Implications for information processing. *Developmental Psychology* 18, 704-713.
- Rothweiler, M. (1989) *Nebensatzwerb im Deutschen. Eine empirische Untersuchung zum Primärspracherwerb*. Dissertation, Universität Tübingen.
- Rovee-Collier, C., Lipsitt, L. P. & Hayne, H. (1998) *Advances in Infancy Research, Vol. 12*. Stamford: Alex Publishing Corporation.
- Saffran, J. R. (2001) Words in a sea of sounds: the output of infant statistical learning. *Cognition* 81, 149-169.
- Saffran, J. R., Aslin, R. N. & Newport, E. L. (1996) Statistical learning by 8-month-old infants. *Science* 274, 1926-1928.
- Saffran, J. R., Johnson, E. K., Aslin, R. N. & Newport, E. L. (1999) Statistical learning of tone sequences by human infants and adults. *Cognition* 70, 27-52.
- Saffran, J. R., Newport, E. L. & Aslin, R. N. (1996) Word segmentation: The role of distributional cues. *Journal of Memory and Language* 35, 606-621.
- Sansavini, A., Bertoncini, J. & Giovanelli, G. (1994) Newborns discriminate stress patterns in phonetically complex Italian words. *Infant Behavior and Development* 17, 924.
- Sansavini, A., Bertoncini, J. & Giovanelli, G. (1997) Newborns discriminate the rhythm of multisyllabic stressed words. *Developmental Psychology* 33, 3-11.
- Santelmann, L. & Jusczyk, P. W. (1998) Sensitivity to discontinuous dependencies in language learners: evidence for limitations in processing space. *Cognition* 69, 105-134.
- Schafer, G. (1998) *Word learning in infancy*. Thesis. University of Oxford.
- Schlesinger, I. M. (1982) *Steps to language: Toward a theory of language acquisition*. Hillsdale: Erlbaum.
- Schlesinger, I. M. (1988) The origin of relational categories. In Y. Levy, I. M. Schlesinger & M. D. S. Braine (eds.) *Categories and processes in language acquisition*. Hillsdale: Erlbaum.
- Scott, D. R. (1982) Duration as a cue to the perception of a phrase boundary. *Journal of the Acoustical Society of America* 71, 996-1007.
- Sebastian-Gallés, N., Dupoux, E., Segui, J. & Mehler, J. (1992) Contrasting syllabic effects in Catalan and Spanish. *Journal of Memory and Language* 31, 18-32.
- Selkirk, E. O. (1984) *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge, MA: MIT-Press.

- Selkirk, E. O. (1996) The prosodic structure of function words. In: J. L. Morgan & K. Demuth (eds.) *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*. Mahwah: Erlbaum.
- Shady, M. & Çerken, L.A. (1999). Grammatical and caregiver cues in early sentence comprehension. *Journal of Child Language* 26, 163-175.
- Shafer, V., Shucard, D., Shucard, J., & Gerken, L. A. (1998) An electrophysiological study of infants' sensitivity of English function morphemes. *Journal of Speech, Language and Hearing Research* 41, 874-886.
- Shi, R., Werker, J. F. & Morgan, J. L. (1999) Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition* 72, B11-B21.
- Shillcock, R. (1990) Lexical hypothesis in Continuous Speech. In G. T. M Altman (ed.) *Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives*. Cambridge, MA: MIT-Press.
- Shipley, E. F., Smith, C. S. & Gleitman, L. R. (1969). A study in the acquisition of language: Free responses to commands. *Language* 45, 322-342.
- Siqueland, E. R. & DeLucia, C. A. (1969) Visual reinforcement of non-nutritive sucking in human infants. *Science* 165, 1144-1146.
- Slobin, D. I. (1973) Cognitive prerequisites for the development of grammar. In C. A. Ferguson & D. I. Slobin (eds.) *Studies of child language development*. New York: Holt, Rinehart & Winston.
- Slobin, D. I. (1985) Crosslinguistic evidence for the language making capacity. In D. I. Slobin (ed.) *The crosslinguistic study of language acquisition, Vol. 2*. Hillsdale: Erlbaum.
- Snow, C. E. (1996) Social perspectives on the emergence of language. In B. MacWhinney (ed.) *The emergence of language*. Mahwah: Erlbaum.
- Spelke, E. (1979) Perceiving bimodally specified events in infancy. *Developmental Psychology* 15, 626-636.
- Spring, D. R. & Dale, P. S. (1977) Discrimination of linguistic stress in early infancy. *Journal of Speech and Hearing Research* 20, 224-232.
- Steinhauer, K., Alter, K. & Friederici, A. D. (1999) Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nature Neuroscience* 2, 191-196.
- Stern, D. N., Spieker, S., Barnett, R. K. & MacKain, K. (1983) The prosody of maternal speech: Infant age- and context-related changes. *Child Language* 10, 1-15.
- Strange, W. & Jenkins, J. J. (1978) Role of linguistic experience in the perception of speech. In: R. D. Walk & H. L. Pick (eds.) *Perception and Experience*. New York: Plenum.
- Streeter, L. A. (1976) Language perception of 2-month old infants shows effects of both innate mechanisms and experience. *Nature* 259, 39-41.
- Streeter, L. A. (1978) Acoustic determinants of phrase boundary perception. *Journal of the Acoustical Society of America* 64, 1582-1592.
- Suomi, K., McQueen, J. M. & Cutler, A. (1997) Vowel harmony and speech segmentation in Finnish. *Journal of Memory and Language* 36, 422-444.
- Swingley, D. & Aslin, R. N. (2000) Spoken word recognition and lexical representation in very young children. *Cognition* 76, 147-166.
- Swingley, D., Pinto, J. P. & Fernald, A. (1998) Assessing the speed and accuracy of word recognition in infants. In C. Rovee-Collier, L. P. Lipsitt & H. Hayne (eds.) *Advances in Infancy Research, Vol. 12*. Stamford: Alex Publishing Corporation.
- Swingley, D., Pinto, J. P. & Fernald, A. (1999) Continuous processing in word recognition at 24 months. *Cognition* 71, 73-108.

- Swoboda, P., Morse, P. A. & Levitt, L. A. (1976) Continuous vowel discrimination in normal and at-risk infants. *Child Development* 47, 459-465.
- Taft, M. & Hambly, G. (1986) Exploring the cohort model of spoken word recognition. *Cognition* 22, 259-282.
- Tallal, P. & Piercy, M. (1973) Developmental aphasia: Impaired rate of non-verbal processing as a function of sensory modality. *Neuropsychologia* 11, 389-398.
- Tallal, P. & Piercy, M. (1974) Developmental aphasia: Rate of auditory processing and selective impairment of consonant perception. *Neuropsychologia* 12, 83-93.
- Tallal, P., & Stark, R. E. (1981) Speech acoustic-cue discrimination abilities of normally developing and language-impaired children. *Journal of the Acoustical Society of America* 69, 568-574.
- Tomasello, M. & Farrar, M. J. (1986) Joint attention and early language. *Child Development* 57, 1454-1463.
- Tomasello, M. (1992) The social bases of language acquisition. *Social Development* 1, 67-87.
- Trehub, S. E. (1973) Infants' sensitivity to vowel and tonal contrasts. *Developmental Psychology* 9, 91-96.
- Trehub, S. E. (1976) The discrimination of foreign speech contrasts by infants and adults. *Child Development* 47, 466-472.
- Treiman, R. & Baron, J. (1981) Segmental analysis ability: Development and relation to reading ability. In: G. E. MacKinnon & T. G. Waller (eds.) *Reading Research: Advances in Theory and Practice, Vol. 3*. New York: Academic Press.
- Treiman, R. & Zukowski, A. (1991) Level of phonological awareness. In: S. A. Brady & D. P. Shankweiler (eds.) *Phonological Processes in Literacy*. Hillsdale, NJ: Erlbaum.
- Umeda, N. (1982) F0-declination is situation dependent. *Journal of Phonetics* 10, 279-290.
- Valian, V. & Coulson, A. S. (1988). Anchor points in language learning: The role of marker frequency. *Journal of Memory and Language* 27, 71-86.
- Valian, V. & Levitt, A. (1996) Prosody and adults' learning of syntactic structure. *Journal of Memory and Language* 35, 497-516.
- Valian, V. (1999) Input and language acquisition. In W. C. Ritchie & T. K. Bhatia (eds.) *Handbook of Child Language Acquisition*. San Diego: Academic Press.
- Van de Weijer, J. (1998) *Language Input for Word Discovery*. Nijmegen, MPI Series in Psycholinguistics
- Van Ooyen, B., Bertocini, J., Sansavini, A., & Mehler, J. (1997) Do weak syllables count for newborns? *Journal of the Acoustical Society of America* 102, 3725-3741.
- Venditti, J. J., Jun S.-A. & Beckman, M. E. (1996) Prosodic cues to syntactic and other linguistic structures in Japanese, Korean, and English. In: J. L. Morgan & K. Demuth (eds.) *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*. Mahwah: Erlbaum.
- Volterra, V., Capirci, O. & Caselli, M. C. (2001) What atypical populations can reveal about language development: The contrast between deafness and Williams syndrome. *Language and Cognitive Processes* 16, 219-239.
- Vroomen, J., Tuomainen, J. & de Gelder, B. (1998) The roles of word stress and vowel harmony in speech segmentation. *Journal of Memory and Language* 38, 133-149.
- Wagner, S. H. & Sakovitz, L. J. (1986) A process analysis of infant visual and cross-modal recognition memory: Implications for an amodal code. In L. P. Lipsitt & C. Rovee-Collier (eds.) *Advances in Infancy Research, Vol. 4*. Norwood, N.J.: Ablex.
- Warren, P., Grabe, E. & Nolan, F. (1995) Prosody, phonology, and parsing in closure ambiguities. *Language and Cognitive Processes* 10, 457-486.

- Weinert, S. (1996) Prosodie – Gedächtnis – Geschwindigkeit: Eine vergleichende Studie zu Sprachverarbeitungsdefiziten dysphasisch-sprachgestörter Kinder. *Sprache & Kognition* 15, 46-69.
- Weissenborn, J. (1991) Functional categories and verb movement in early German. The acquisition of German reconsidered. In: M. Rothweiler (ed.) *Spracherwerb und Grammatik: Linguistische Untersuchungen zum Erwerb von Syntax und Morphologie*. Linguistische Berichte, Sonderband 3/1990.
- Weissenborn, J. (1994) Constraining the child's grammar: Local wellformedness in the development of verb movement in German and French. In: M. Suner, B. Lust & J. Whitman (eds.) *Syntactic Theory and Language Acquisition: Crosslinguistic Perspectives. Vol. 1: Phrase Structure*. Hillsdale, NJ: Erlbaum.
- Weissenborn, J. (2000) Der Erwerb von Morphologie und Syntax. In H. Grimm (ed.) *Enzyklopädie der Psychologie, Band 3: Sprachentwicklung*. Göttingen: Hogrefe.
- Weissenborn, J. & Höhle, B. (2001) Introduction. In J. Weissenborn & B. Höhle (eds.) *Approaches to Bootstrapping: Phonological, lexical, syntactic, and neurophysiological aspects of early language acquisition, Vol. 1*. Amsterdam: Benjamins.
- Weissenborn, J., Goodluck, H. & Roeper, T. (1992) Old and new problems in the study of language acquisition. In J. Weissenborn, H. Goodluck & T. Roeper (eds.) *Theoretical Issues in Language Acquisition*. Hillsdale: Erlbaum.
- Weissenborn, J., Höhle, B., Kiefer, D. & Cavar, D. (1998) Children's sensitivity to word-order violations in German: Evidence for very early parameter-setting. In A. Greenhill, M. Hughes, H. Littlefield, & H. Walsh (eds.) *Proceedings of the 22nd Annual Boston Conference on Language Development*. Somerville: Cascadilla Press.
- Wenk, B. & Wiolland, F. (1982) Is French really syllable-timed? *Journal of Phonetics* 10, 193-216.
- Werker, J. F. (1991) The ontogeny of speech perception. In I. G. Mattingly & M. Studdert-Kennedy (eds.) *Speech and perception and modularity*. New Jersey: Erlbaum.
- Werker, J. & Tees, R. (1984) Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development* 7, 49-63.
- Werker, J. F. & Lalonde, C. E. (1988) Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology* 24, 672-683.
- Werker, J. F., Gilbert, J. H., Humphrey, K. & Tees, R. C. (1981) Developmental aspects of cross-language speech perception. *Child Development* 52, 349-355.
- Werker, J. F., Lloyd, V. L., Pegg, J. E. & Polka, L. (1996) Putting the baby in the bootstraps: Toward a more complete understanding of the role of the input in infant speech processing. In J. L. Morgan & K. D. Demuth (eds.) *Signal to syntax: Bootstrapping from speech to syntax in early language acquisition*. Hillsdale, N.J.: Erlbaum.
- Werker, J. F., Lloyd, V. L., Pegg, J. E. & Polka, L. (1996) Putting the baby in the bootstraps: Toward a more complete understanding of the role of the input in infant speech processing. In: J. L. Morgan & K. Demuth (eds.) *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*. Mahwah: Erlbaum.
- Whalen, D. H., Levitt, A. G. & Wang, Q. (1991) Intonational differences between the reduplicative babbling of French- and English-learning infants. *Journal of Child Language* 18, 501-516.
- Wiese, R. (1996) *The Phonology of German*. Oxford: Clarendon Press.
- Zwitserslood, P. (1989) The locus of the effects of sentential-semantic context in spoken-word processing. *Cognition* 32, 25-64.
- Zwitserslood, P., Schriefers, H., Lahiri, A. & van Donselaar, W. (1993) The role of syllables in the perception of spoken Dutch. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19, 260-271.