

1 Aspekte der Sprache und theoretische Zugänge

1.1 Komponenten der Sprache

Wenn wir sprechen, machen wir uns meist keine Gedanken darüber, welcher komplizierter Vorgang das ist, der im Allgemeinen automatisch und unbewusst abläuft. Die Worte „entstehen in unserem Kopf“ und „kommen aus unserem Mund“. Sprache, die wir produzieren oder die wir hören, kann dabei unter verschiedenen Aspekten betrachtet werden, die alle zusammen Sprache konstituieren. Gesprochene Sprache besteht zunächst aus Lauten. Im Mund-Nasen-Rachenraum, dem Ansatzrohr, wird mit der Zunge, den Lippen oder den Zähnen der im Kehlkopf erzeugte Ton so verändert, dass einige Dutzend Konsonanten und Vokale entstehen. Die so gebildeten Laute werden in den verschiedenen Sprachen nach unterschiedlichen Regeln kombiniert, die zu Wörtern verbunden werden (Lieberman & Liberman, 1992). Individuelle Sprachen nutzen dabei nur einen Bruchteil der möglichen Sprachlaute und Kombinationen (Barret, 1999). In diesen Sprachlauten und ihren Kombinationen wird Bedeutung kodiert. *Phoneme*, also Laute, sind die kleinsten bedeutungsunterscheidenden Einheiten. Sie haben für sich noch keine Bedeutung, lassen aber beispielsweise aus „Tanne“ „Kanne“ werden.

Die *Phonetik* als Teil der Lehre von den Lauten, betrifft ihre physikalische Struktur und beschreibt, wie sie mit den menschlichen Sprechwerkzeugen hervorgebracht werden. Unter *artikulatorischer Perspektive* wird in der Phonetik die Produktion der Laute betrachtet. Wenn im Kehlkopf ein Ton erzeugt und der Luftstrom nicht blockiert wird, entstehen *Vokale*, die ihren Klang durch die unterschiedliche Öffnung des Mundraumes erhalten. Wenn der Luftstrom auf unterschiedliche Weise bei der Passage durch den Rachen-Mund-Raum blockiert wird, entstehen *Konsonanten*. Man kann drei Gruppen von phonetischen Merkmalen unterscheiden, nach denen sich alle Laute klassifizieren lassen, die *Sonorität* (stimmlos/stimmhaft), die *Artikulationsstelle* (die Stellen, an denen der Luftstrom blockiert wird) und die *Artikulationsart* (welcher Laut entsteht), z. B. Plosive wie /p/, /b/, /t/, /d/, /k/ oder /g/, bei denen der Luftstrom zunächst vollkommen blockiert und dann „explosionsartig“ freigesetzt wird, Frikative (Reibelaute) wie /f/, /v/ oder die verschiedenen s- und Zisch-Laute, Nasale wie /n/ und /m/, bei denen die Luft durch die Nase entweicht, Laterale wie /l/, bei denen der Luftstrom seitlich austritt, und den Vibranten /r/ und Kombinations-

nen, wie beispielsweise Affrikative, bei denen ein eigentlicher Plosivlaut nicht vollständig aufgelöst wird, sondern in einen Reibelaut übergeht, wie bei /pf/. Unter der *akustischen Perspektive* werden die Merkmale der Schallsignale als Muster von Frequenzen und Intensitäten im zeitlichen Verlauf betrachtet. Sie lassen sich in einem Schallspektrum darstellen.



Abb. 1.1: Schallspektrum einzeln und deutlich gesprochener Wörter (aus Goldstein, E. B. (1997). *Sensation and Perception*. Washington: Itps Thomson Learning)

Unter *auditiver Perspektive* werden die psychischen Vorgänge bei der zentralen Verarbeitung der Sprachlaute betrachtet. Dabei gibt es keine Eins-zu-eins-Zuordnung zwischen den physikalischen Merkmalen und den wahrgenommenen Lauten. Sprachsignale sind flüchtig, müssen schnell erfasst und trotz unterschiedlicher Sprecher, nicht „hörbarer“ Wortgrenzen im Sprachstrom und anderer Variationen als Einheiten erkannt werden.

Abbildung 1.2 zeigt, wie sich das Schallspektrum verändert, wenn der Satz „What are you doing?“ in der Alltagssprache gesprochen wird. Pro Sekunde müssen bei durchschnittlicher Sprechgeschwindigkeit ca. 30 Phoneme erkannt werden, während man gleichzeitig Wortbedeutungen, Satzstrukturen und Satzbedeutungen analysiert (Penner, 2000). Allein die phonetischen Kompetenzen und der Erwerb des Lautbestandes der Sprache und der in der Umgebungssprache erlaubten Kombinationsmöglichkeiten sind Leistungen, die ein Kind in seinen ersten Lebensjahren vollbringt, deren Komplexität dem naiven Betrachter kaum präsent ist. Die *Phonologie* betrachtet die Lautsysteme einer Sprache, ihre Repräsentationen, also die Zuordnung der bei verschiedenen Personen unterschiedlich gebildeten Laute zu einem Phonem als abstrakter Lauteinheit.

Laute und ihre Kombinationen werden zu Wörtern verbunden und repräsentieren Bedeutungen, dies betrifft die *Semantik*. Wörter stehen für Begriffe, Konzepte und Beziehungen und verbinden damit sprachliche mit kognitiven Fähigkeiten. Die kleinsten bedeutungstragenden Einheiten sind die *Morpheme*, die

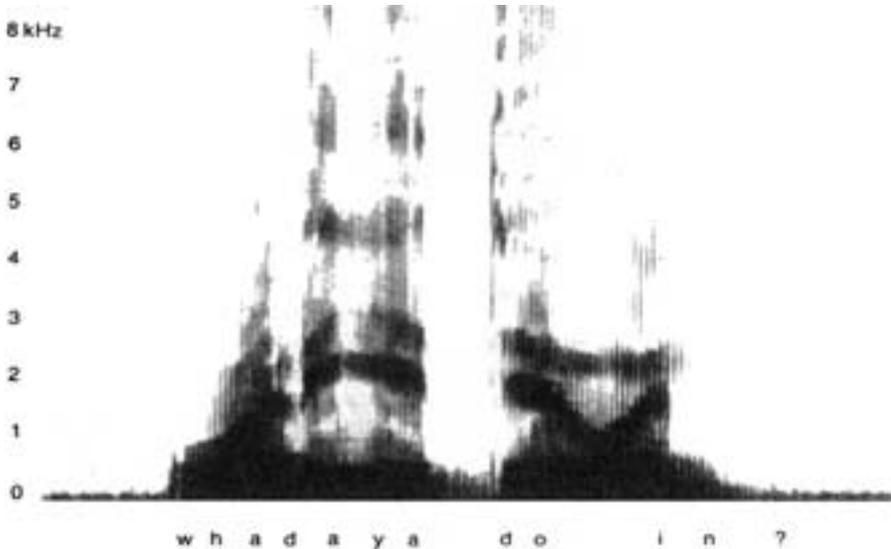


Abb. 1.2: In Alltagssprache gesprochener Satz (aus Goldstein, E. B. (1997). *Sensation and Perception*. Washington: Itps Thomson Learning)

frei sein können, wie „jung“, „Haus“ oder „Spiel“ oder gebunden wie „heit“, „keit“, „un“ oder „ver“. Wenn sie an ein freies Morphem geheftet werden, machen sie beispielsweise aus einem Verb ein Adjektiv (vergessen – vergesslich) und daraus ein Substantiv (vergesslich – Vergesslichkeit). Wie Morpheme in einer Sprache kombiniert werden dürfen, beschreibt die *Morphologie*. Morpheme können *derivativ* (Ableitungsmorpheme) sein, wenn sie die Wortklassen verändern (heiter – Heiter-keit), oder *flexiv* (Beugungsmorphem), wenn sie Wörter beugen (ich spring-e – du spring-st). Die *Syntax* beschreibt die Kombinationsregeln von Wörtern in einem Satz, durch die Sätze, die aus den gleichen Wörtern bestehen, unterschiedliche Bedeutungen erhalten können. Beispielsweise bedeutet „Hans liebt Grete“ tragischerweise nicht dasselbe wie „Grete liebt Hans“. Allerdings bedeutet „Grete wird von Hans geliebt“ dasselbe wie „Hans liebt Grete“, ein Ausdruck grammatischer Regeln, die angeben, wie Wörter und Morpheme in einer bestimmten Sprache kombiniert, organisiert und in eine Reihenfolge gebracht werden dürfen.

Schließlich muss das Kind noch lernen, wie Sprache in unterschiedlichen Kontexten gebraucht (*kommunikative Funktionen*), wie Rede und Gegenrede gestaltet wird (*Konversation*) und wie man seine Rede zusammenhängend aufbaut und auf den Informationsstand und die Perspektive des Partners abstimmt (*Diskurs*). Alle drei Funktionen gehören in den Bereich der *Pragmatik*. In der Pragmatik werden damit die linguistischen Kompetenzen mit sozialen Fähigkeiten verbunden.

Als weitere, suprasegmentale, also über den untergliederbaren Einheiten liegende Komponente, gilt die *Prosodie*. Damit wird die Betonung und Gliederung

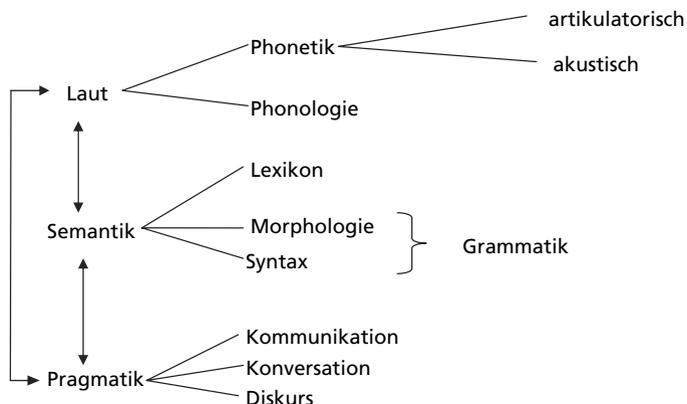


Abb. 1.3: Komponenten der Sprache (nach Barret, 1999, S. 8)

der Sprache bezeichnet, die sich durch Tonhöhe, Lautheit, Länge der Sprachlaute und Pausengebung darstellt (Grimm, 1999; Grimm & Weinert, 2002). Durch die typischen Betonungsmuster können Sprachen voneinander unterschieden werden, auch wenn man sie nicht beherrscht. Wir „hören“, ob es sich um eine slawische, eine romanische, eine arabische oder eine asiatische Sprache handelt. Innerhalb einer Sprache gibt die Prosodie Hinweise, wie eine Äußerung zu verstehen ist. „Ich will *den* Ball“ (und keinen anderen), „Ich will den *Ball*“ (und nichts anderes), „Ich *will* den Ball“ (du musst ihn mir geben!), „*Ich* will den Ball“ (du kriegst ihn nicht!) wird je nach Betonung anders verstanden.

1.2 Das Sprachlernproblem

Lange hat man den Beginn der Sprachentwicklung mit dem Auftauchen des ersten Wortes gegen Ende des ersten Lebensjahres gleichgesetzt. Heute weiß man, dass Sprachentwicklung bereits vor der Geburt des Kindes beginnt und Einjährige schon wichtige Schritte auf dem Weg in die Sprache unbemerkt vollzogen haben. Wie sie das tun, warum sie das im Allgemeinen so perfekt schaffen, welche Lernmechanismen sie zur Verfügung haben, beschäftigt Psychologen, Psycholinguisten und Linguisten seit Jahren, ohne dass sie bisher zu einem allgemein akzeptierten Ergebnis gekommen wären. Zwar wird niemand bestreiten, dass „normale Kinder in normaler Umgebung Sprache erlernen“ (Hoff-Ginsberg, 2000, S. 463), genauso wenig wie die Tatsache, dass es bei „normalen Kindern in normaler Umgebung“ erhebliche Unterschiede in der Geschwindigkeit und im Verlauf der Sprachentwicklung gibt. Welche sprachspezifischen und/oder welche allgemeinen Fähigkeiten das Kind benötigt, erlernt oder mitbringt und welche Rolle der sprachliche Input dabei spielt, wird je nach der Theorie zugrundeliegender Annahme unterschiedlich gesehen.

Grundsätzlich kann man drei große Gruppen von Spracherwerbtheorien unterscheiden, die als *inside-out*, *outside-in* und die radikale Mitte gekennzeichnet werden können (Hennon, Hirsh-Pasek & Golinkoff, 2000; Hirsh-Pasek & Golinkoff, 1996).

- *Inside-out-Theorien* beschreiben den Spracherwerb überwiegend als Ausdruck einer angeborenen, bereichsspezifischen Prädisposition. Ein Minimum an sprachlichem Input, der nicht einmal vollkommen sein muss, triggert das bereits vorhandene Sprachlernprogramm (Chomsky, 1965, 1981; Pinker, 1996, 2003). Sprache wird gedacht als ein Modul oder eine Gruppe von Modulen. Module sind voneinander unabhängige Einheiten (*encapsulation*), die einen bestimmten Input mit nur diesem Modul eigenen Prozessen verarbeiten (*domain specific*). Sie werden nicht oder kaum von anderen Modulen und Prozessen beeinflusst und sind zur unwillkürlichen Verarbeitung des Inputs zwingend nötig (*mandatory*) (Fodor, 1983; Temple, 1997). *Inside-out*-Vertreter gehen also von angeborenen, sprachspezifischen Mechanismen aus, die die Teilprozesse der Sprache verarbeiten.
- *Outside-in-Theorien* dagegen sehen den Hauptmotor der Sprachentwicklung in den sozialen Interaktionen in der sprechenden Umgebung (Bruner, 1983; Tomasello, 2003a). Das Kind abstrahiert nach dieser Grundannahme die sprachlichen Regeln seiner Umwelt und folgt dabei allgemeinen kognitiven Mechanismen. Es kommt zur Sprache, indem es im Gebrauch (*usage-based*) Intentionen erkennt und Muster anhand der Auftretenshäufigkeiten herausfiltert (L. Bloom, 1998; Nelson, 1996, 1999; Tomasello, 2003a, 2003b).

Beide Auffassungen können als Extreme eines Kontinuums angesehen werden (Hirsh-Pasek & Golinkoff, 1996) und beide Auffassungen können die Frage, wie Sprache gelernt wird, nicht vollständig erklären. *Outside-in*-Vertreter können nicht ausreichend erklären, warum Kinder auf dem Weg in die Sprache viele mögliche induktive Schlüsse nicht ziehen, sondern relativ zielsicher die „richtigen“, während *Inside-out*-Vertreter, die von einer sprachspezifischen Vorausstattung ausgehen, zunächst erklären müssen, wie das Kind die relevanten Einheiten aus dem Strom der gesprochenen Sprache erkennen kann (Tracy, 2000). Die Theorie muss auch erklären können, wie die abstrakte, als angeboren angenommene, universelle Grammatik an die verschiedenen Zielsprachen gebunden wird (*linking problem*), und warum sich auf der Welt so viele sehr unterschiedliche Sprachen gebildet haben, wenn alle Menschen mit derselben sprachspezifischen Ausstattung geboren werden. Im Verlauf der Sprachentwicklung eines Kindes lassen sich qualitative Veränderungen feststellen, die erklärt werden müssen, wenn die universelle Grammatik immer dieselbe ist (*continuity problem*) (Tomasello, 2003a).

- Die *rational-konstruktivistische* oder „*radikale Mitte*“ (Grimm, 1999; Grimm & Weinert, 2002; Hennon et al., 2000; Hirsh-Pasek & Golinkoff, 1996) vereint beide Auffassungen. Das Kind bringt Voreinstellungen zum Erwerb der Sprache mit und konstruiert sie in Wechselwirkung mit den Angeboten seiner Umgebung.

Die Geschwindigkeit und der frühe Zeitpunkt, zu dem ein Kind Sprache erwirbt, sprechen dafür, dass es eine biologisch determinierte Bereitschaft auf Sprache zu achten, sprachrelevante Reize bevorzugt zu verarbeiten und sprachliche Regeln zu erkennen, mitbringt. In sozialen Interaktionen, in der Ausbildung von Routinen und Mustern und in kommunikativen Anforderungen, kann das Kind das sprachliche Angebot der Bezugspersonen nutzen, um aus den relevanten Informationen sein Sprachwissen zu konstruieren. Die Bezugspersonen zeigen intuitiv ein sprachliches und kommunikatives Verhalten, das dem Kind den „Einstieg“ in die Sprache erleichtert, indem sie zum Kind in kurzen, grammatikalisch korrekten Sätzen sprechen und dabei die sinntragenden Wörter besonders betonen. Aber selbst dieses, oft als universell angenommene Verhalten der Eltern, die kindgerichtete Sprache (Ammensprache, Babytalk) gilt nicht für alle Kulturen (Lieven, 1994; Rüter, 2004; Schieffelin, 1985). Je nach Einstellung zum Kind als ein grundsätzlich verstehendes oder grundsätzlich (noch nicht) verstehendes Wesen unterscheidet sich der sprachliche Umgang mit den Kindern, so dass zwei verschiedene Lernstrategien zu beobachten sind. Während in manchen Kulturen, die davon ausgehen, dass Kinder eine gewisse Reife erwerben müssen, um dann „von selbst“ zu sprechen, ein „Sprich erst und analysiere dann!“ gilt, nehmen Eltern unserer westlichen Kulturen an, dass Kinder Sprache von Anfang an aufnehmen. Sie verhelfen ihnen mit ihrer kindgerichteten Sprache dazu, zunächst einzelne Einheiten und Komponenten der Sprache zu analysieren, um sich auf dieser Grundlage die Sprache konstruktiv anzueignen, also mit der Strategie „Analysiere erst und sprich dann!“ vorgehen zu können (Hoff-Ginsberg, 2000).

Das Sprachlernproblem ist vermutlich am ehesten unter einer entwicklungspsychologischen Perspektive zu verstehen, bei der weder die Anlage noch die Umwelt als entscheidender Faktor zu sehen sind, sondern das Zusammenspiel beider (L. Bloom, 1998; Gopnik & Nazzi, 2003; Grimm & Wilde, 1998; Hennon et al., 2000; Hirsh-Pasek & Golinkoff, 1996). Die inneren Voraussetzungen des Kindes und die lernbegünstigende Umgebung müssen passen (Grimm & Weinert, 2002). Wörter und Grammatik sind nicht von Anfang an vorhanden, sondern entstehen dann, wenn die biologischen Voreinstellungen mit der sozialen Umwelt interagieren.

Übungsfragen zu Kapitel 1

1. Wie können Sprachlaute klassifiziert werden?
2. Welche Komponenten konstituieren die menschliche Sprache?
3. Was spricht für Outside-in-Theorien?
4. Was spricht für Inside-out-Theorien?
5. Was spricht für die „radikale Mitte“?

2 Sprachentwicklung

2.1 Spracherwerb im Säuglings- und Kleinkindalter

2.1.1 Prosodie und Phonologie

Bevor das Kind sein erstes Wort spricht, hat es schon sehr viel über Sprache gelernt. Die Methoden der Säuglingsforschung sind in den letzten Jahrzehnten sehr verfeinert worden und erlauben, Wahrnehmungs- und Denkleistungen des Säuglings zu erkennen, auch wenn er sich noch nicht äußern und motorisch noch kaum gezielt reagieren kann. Durch moderne und einfallreiche Projektions- und Computertechniken ist es möglich, die Erfassung der kognitiven Leistungen von den motorischen abzukoppeln, so dass bereits für Neugeborene differenzierte Wahrnehmungsleistungen nachgewiesen werden können (zusammenfassend s. z. B. Goswami, 1998; Jusczyk, Houston & Goodman, 1998; Mandler, 1998).

Drei Untersuchungsparadigmen, die prinzipiell auf dem Wechselspiel von Habituation, also Gewöhnung, und Dishabituation, also Orientierungsreaktion, beruhen, können unterschieden werden. Im *Habituationparadigma* wird einem Baby so lange ein bestimmter Reiz, ein Bild, eine Stimme, eine Lautfolge, dargeboten, bis keine Reaktionen auf den Reiz mehr zu beobachten sind, es sich abwendet, eindöst, missmutig wird oder andere Anzeichen von Desinteresse zeigt. Wenn man dann die Reizkonfiguration ändert und das Baby wieder aufmerksam schaut oder hört, also eine Dishabituation ausgelöst wird, ist damit deutlich, dass es einen Unterschied erkannt hat. Durch Manipulation des Untersuchungsmaterials kann man herausfinden, welche perzeptuellen Unterschiede das Baby bereits erkennen kann. Habituation und Dishabituation lässt sich auch „objektiver“ anhand der Veränderung der Herzschlagfrequenz schon beim Ungeborenen oder der Saugrate beim Säugling feststellen.

Beim *Präferenzparadigma* werden dem Baby zwei Reize geboten, beispielsweise zwei Stimmen, zwei Gesichter, zwei Gegenstände. Wenn das Baby eins davon intensiver, d. h. länger (Fixationsrate) betrachtet als das andere, wird das als Bevorzugung interpretiert. Dabei kommt es natürlich darauf an, dass die Versuchsbedingungen so variiert werden, dass keine Reihenfolgeeffekte auftreten. Präferenzen können auch festgestellt werden, wenn speziell präparierte Sauger mit

Dia-, Tonband- oder Videogeräten gekoppelt werden, so dass der Säugling mit seiner Nuckelfrequenz die Geräte einschalten und sich so das bevorzugte „Programm“ selbst verschaffen kann. Mithilfe des Präferenzparadigmas kann auch die Kopplung zweier Modalitäten untersucht werden. Wenn einem vier Monate alten Baby zwei Videoclips gleichzeitig gezeigt werden, wobei auf dem einen eine Frau „Kuckuck – da“ (ihr Gesicht mit den Worten „Kuckuck“ verdeckt und bei „da“ freudig wieder hervorkommt) spielt und auf dem anderen mit einem Stock auf einen Holzblock geschlagen wird, schaut das Baby bevorzugt zu dem Clip, der mit der Tonspur übereinstimmt (Spelke, 1976).

Beim *Überraschungsparadigma* wird dem Kind ein „unmögliches Ereignis“ gezeigt. Beispielsweise wird es darauf habituiert, dass ein Hase hinter einem Schirm verschwindet und am anderen Ende wieder herauskommt. In der Dishabituationphase wird im Schirm ein Rechteck ausgeschnitten, so dass ein großer Hase im Ausschnitt zu sehen sein müsste, wenn er hinter dem Schirm vorbeigeht, ein kleiner jedoch nicht. Babys im Alter von fünf Monaten reagieren irritiert, wenn der große Hase im Ausschnitt nicht zu sehen ist, aber nicht, wenn der kleine hinter dem Schirm verschwindet (Baillargeon & Graber, 1987). Gerade Experimente mit Überraschungen, unmöglichen Ereignissen, die sich meist nur mit Tricktechniken herstellen lassen, erlauben Schlüsse auf kognitive Prozesse des jungen Säuglings, da derartige Ereignisse in der physikalischen Welt nicht vorkommen, sie also nicht einfach nur wiedererkannt sein können.

Bei der Untersuchung der Sprachwahrnehmung von Säuglingen werden als Abwandlungen dieser Paradigmen hauptsächlich die konditionierte Kopfhinwendung (conditioned headturn, CHT), die visuelle Habituation (VH) und die erhöhte Saugfrequenz (high-amplitude sucking, HAS) eingesetzt (Werker et al., 1998).

Bei der *konditionierten Kopfhinwendung* (CHT) wird das Baby darauf trainiert, dass sofort nach jedem neuen Stimulus seitlich vom Kind ein lustiges, buntes Spielzeug (z. B. ein – lautlos – trommelnder Stoffaffe) in Gang gesetzt wird. Beispielsweise hört das Baby über einen gewissen Zeitraum den Laut /ba/, der dann in /da/ gewechselt wird, während gleichzeitig der Affe zu trommeln beginnt, so dass das Baby den Zusammenhang zwischen Reizpräsentation und Bewegung des Spielzeugs lernt (erste Konditionierungsphase). In einer zweiten Konditionierungsphase wird eine immer längere Pause zwischen dem neuen Reiz und der Bewegung des Spielzeugs eingeführt, so dass das Baby bei der Präsentation des neuen Reizes bereits in Erwartung der Aktivierung des Spielzeugs seinen Kopf dreht. In der Versuchsphase werden neue Reize und gleichbleibende Kontrollreize in zufälliger Reihenfolge geboten. Wenn das Baby den Kopf dreht, obwohl kein neuer Reiz kam, bleibt der Affe ruhig. Ebenso sieht es ihn nicht, wenn es den Kopf nicht gedreht hat, obwohl ein neuer Reiz zu hören war. Das Hinwenden des Kopfes in der Erwartung, dass der Affe zu trommeln beginnt, zeigt an, dass das Baby die Reize diskriminiert hat. Natürlich müssen alle anderen Personen im Raum mit über Kopfhörer eingespielter Musik „veräubt“ werden, um dem Kind nicht verdeckte Hinweise auf die Reaktion zu ge-

ben. Stimulus (Ton) und Verstärker (Spielzeug) sind voneinander unabhängig. Das Baby verschafft sich durch das Hinwenden des Kopfes den Verstärker aktiv selbst, so dass besser entschieden werden kann, ob es Unterschiede nicht wahrnimmt oder einfach nur desinteressiert oder unkooperativ ist.

Dies kann bei der *visuellen Habituation (VH-Methode)* ein Problem sein. Hier wird die Aufmerksamkeit des Babys auf einen Bildschirm gelenkt. Während es dort ein Bild betrachtet, ertönt ein Stimulus. Nach einer Weile hat es sich daran gewöhnt, es hat habituiert und betrachtet das Bild nicht mehr. Wenn dann der Stimulus geändert wird, schaut das Baby wieder hin. Allerdings sind nur positive Reaktionen interpretierbar. Wenn das Kind nicht reagiert, könnte es auch den Reiz nicht interessant genug finden, so dass es keine Reaktion zeigt, obwohl es den Unterschied erkannt hat.

Die Messung der *erhöhten Saugrate (HAS-Methode)* eignet sich für sehr junge Säuglinge bis zu vier Monaten. Mit entsprechend präparierten Schnullern kann die Saugrate des Babys aufgezeichnet werden und mit dem Wechsel von Reizen wie Bildern oder Videoclips verbunden werden, so dass beispielsweise ein neues Bild erscheint, wenn eine bestimmte Saugfrequenz erreicht ist. Neugeborene saugen zwar stärker, wenn ihnen neue Reize geboten werden. Sie erkennen aber vermutlich die Kontingenz zwischen Nuckeln und Auftauchen des Reizes noch nicht. Zwei bis vier Monate alte Kinder „erkennen“ den Zusammenhang und können sich durch eine erhöhte Saugrate ein bevorzugtes Bild aktiv verschaffen (Werker et al., 1998). Mit diesen Untersuchungsmethoden können vor allem die auditiven Wahrnehmungsleistungen des Säuglings festgestellt werden, die eine Grundvoraussetzung sind, „auf natürlichem Weg“ sprechen zu lernen.

Bereits ab der 27./28. Schwangerschaftswoche kann der Fötus hören (Hennon et al., 2000; Lecanuet, 1998). Er hört allerdings durch den dämpfenden Filter des Uterus hauptsächlich die prosodischen Merkmale, an denen er später die Stimme seiner Mutter erkennt (Fifer & Moon, 1989; Höhle, 2005). Er kann vertraute kurze „Geschichten“ wiedererkennen (DeCasper et al., 1994) und bevorzugt sie (DeCasper & Spence, 1986). Er kann die prosodischen Merkmale seiner Muttersprache wahrnehmen und nach der Geburt wiedererkennen (DeCasper et al., 1994). Er kann Sprachlaute von Geräuschen unterscheiden und sogar Abfolgen von Silben und Betonungsmuster differenzieren (Lecanuet, Granier-Deferre & Busnel, 1995). Das Neugeborene kann verschiedene Sprachen anhand prosodischer Merkmale voneinander unterscheiden und bevorzugt die Muttersprache (Moon, Cooper & Fifer, 1993). Bereits im Mutterleib beginnt also eine sprachliche Entwicklung, durch die das Neugeborene mit einer Ausstattung zur Welt kommt, die es besonders auf sprachliche Reize achten lässt (Lecanuet, 1998). Inwieweit es sich dabei um sprachspezifische Prädispositionen handelt (Jusczyk, 2001, 2003; Jusczyk et al., 1998) oder um allgemeine, kognitive Mechanismen, die besonders gut auf sprachliche Reize angewendet werden können, so dass sie sich schließlich immer sprachspezifischer entwickeln, ist bislang ungeklärt.

Gesprochene Sprache ist im ständigen Fluss, Wortgrenzen sind nicht hörbar. Pausen liegen häufiger zwischen Silben als zwischen Wörtern. Laute verändern

sich, je nach ihrer Position im Wort. Vorausgehende Laute bereiten die folgenden in „*Koartikulation*“ vor und verändern damit ihre Bildung. Bei dem Wort „Boot“ wird das „B“ mit geschürzten Lippen gebildet, um das „o“ vorzubereiten, während bei „Bad“ die Lippen breit geschlossen bleiben für das folgende „a“. Wenn Wörter in Silben segmentiert gesprochen werden (Ma-ma), lassen sie sich immer noch als das gemeinte Wort identifizieren, nicht aber, wenn die Laute einzeln gesprochen werden (M-a-m-a). Oft wird daher nicht das Phonem als natürliche Einheit gesehen, sondern die Silbe (M.J. Adams, 1991; L. Bloom, 1998; Eimas, 1985; Gleitman et al., 1989; Goldstein, 1997; Goodman, 1984; Liberman & Liberman, 1992; Marcel, 1980; Penner, 2000; Segui, Frauenfelder & Hallé, 2001).

Um letztlich Sprache verstehen zu können, muss der Lautstrom in die für das Verständnis relevanten Einheiten zergliedert werden. Einen ersten Anhaltspunkt für diese Segmentierung bieten dem Säugling die prosodischen Merkmale seiner Umgebungssprache. Dabei bevorzugt er die natürlichen Muster seiner Muttersprache, die in der an das kleine Kind gerichteten Sprache, der „*Ammensprache*“ besonders betont werden (Grimm & Wilde, 1998; Hennon et al., 2000; Papoušek, 1994; Papoušek & Haekel, 1987). Erwachsene und auch ältere Kinder sprechen in einer hohen Stimmlage in kurzen, korrekten und grammatisch einfachen Sätzen zu Säuglingen und Kleinkindern, dehnen die betonten Wörter, modulieren ihre Stimme stärker und bieten ihnen damit genau die Merkmale deutlich, die sie zur Segmentierung nutzen. Allerdings ist zweifelhaft, ob die Erwachsenen das in lehrender Absicht tun oder nicht eher im Bemühen um Kommunikation mit einem nicht-kompetenten Sprecher, um sein Verständnis zu verbessern und seine Aufmerksamkeit aufrechtzuerhalten (Pine, 1994).

Ein neugeborenes Baby kann die meisten menschlichen Sprachlaute voneinander unterscheiden (*Phonemdiskriminanz*). Die auditiven Erfahrungen mit der Umgebungssprache spielen dabei offensichtlich noch eine untergeordnete Rolle (Höhle, 2004). Es ist als „*Sprachgeneralist*“ in der Lage, jede Sprache, die in seiner Umgebung gesprochen wird, zu lernen. Um seine Muttersprache zu erwerben, muss es jedoch eine weitere Anforderung erfüllen: Es muss Laute, die sich akustisch unterscheiden, als gleich auffassen können (*Phonemkonstanz*) (Bishop, 1997; Kuhl, 1985).

Menschliche Sprachlaute befinden sich auf einem Kontinuum, so dass eine unendliche Zahl von Sprachlauten möglich wäre, einzelne Sprachen verwenden jedoch nur einen begrenzten Teil davon. Die Laute werden, ob nun eine Frau, ein Kind oder ein Mann spricht und welche individuellen Unterschiede in den Sprechwerkzeugen bestehen, wegen der Koartikulation so unterschiedlich gebildet, dass sie im Schallspektrum nicht als gleich identifiziert werden können. Das Baby muss also lernen, die Konstanz trotz dieser Variationen zu erkennen. Das gelingt, indem „*Prototypen*“ von Lauten gebildet werden, „*beste*“ phonologische Repräsentationen, an die alle gehörten Laute angeglichen werden (Gopnik, Kuhl & Meltzoff, 2001). Laute werden zu Kategorien zusammengefasst, wobei gewisse Abweichungen geduldet und der Lautstrom nicht mehr als sich kontinuierlich verändernd, sondern in diskreten Klassen von Lauten zusammengefasst, wahrgenommen wird.