

# Sprachwahrnehmung im Säuglingsalter

Nicole Altvater-Mackensen 

Anglistik I / Psycholinguistik, Philosophische Fakultät, Universität Mannheim, Deutschland

**Zusammenfassung:** Der Artikel gibt einen Überblick über die Entwicklung der Sprachwahrnehmung im ersten Lebensjahr. Er zeichnet den Entwicklungsverlauf von den frühesten Hörerfahrungen bis zum ersten gesprochenen Wort nach und geht darauf ein: (i) wie vorgeburtliche Erfahrungen das Hören bei Neugeborenen prägen; (ii) wie die Sprachwahrnehmung sich im Säuglingsalter verändert und auf die Umgebungssprache einstellt; (iii) welche Faktoren die Veränderung in der Sprachwahrnehmung beeinflussen; (iv) wie theoretische Modelle diese Veränderungen erklären; und (v) welche Rolle die perzeptuelle Entwicklung im ersten Lebensjahr für die weitere Sprachentwicklung spielt.

**Schlüsselwörter:** Sprachwahrnehmung, sprachliche Entwicklung, perceptual attunement, Säuglingsalter

## Speech Perception in Infancy

**Abstract:** Before they speak their first word, infants reach important milestones in their language development. In particular, their speech perception becomes attuned to the ambient language, first to its prosodic and then to its segmental properties. This article reviews developmental changes in speech perception across the first year of life. It begins with the earliest hearing experiences and then discusses how prenatal exposure shapes listening preferences in newborns, thereby providing a basis for further development. It furthermore sketches how infants attune their perception of sound contrasts to native phonological categories, both auditorily and audio-visually, at the cost of reduced sensitivity to nonnative contrasts. The second part of the article considers the factors that influence the infant's perceptual development, including quantitative and qualitative features of the language input and influences of emerging articulatory abilities. It further details different theories on the processes and mechanisms underlying perceptual attunement. It discusses how bottom-up and top-down information modulates learning, such as statistical and lexical cues to phonemic contrasts. The review concludes with an outline of how perceptual development in the first year of life links to further language development.

**Keywords:** speech perception, language development, perceptual attunement, infancy

Kinder scheinen Sprache mühelos zu lernen. Bereits um den ersten Geburtstag herum sprechen Kleinkinder ihr erstes Wort (Kauschke, 2012), mit zwei Jahren kennen sie mehrere Hundert Wörter (Szagun, 2019), und mit vier Jahren erzählen sie bereits Geschichten (Hellrung, 2019). Diese produktiven Meilensteine sind der Beleg sprachlicher Entwicklung, die auf Fähigkeiten der Sprachwahrnehmung aufbaut. Der vorliegende Artikel fasst die Entwicklung der Sprachwahrnehmung im ersten Lebensjahr zusammen und zeichnet den Entwicklungspfad von den frühesten Hörerfahrungen bis zum ersten Wort nach. Der Überblick ist konzeptionell angelegt und beschreibt beispielhaft Entwicklungsverläufe anhand wegweisender Studien, ergänzt durch Verweise auf aktuelle Überblicksarbeiten und Meta-Analysen. Beginnend bei der pränatalen Entwicklung wird beschrieben wie sich die Sprachwahrnehmung im Säuglingsalter verändert und an die Umgebungssprache anpasst. Im Anschluss wird darauf eingegangen, welche Faktoren diese Veränderung beeinflussen und wie verschiedene theoretische Modelle sie

erklären. Abschließend werden die Auswirkungen der Entwicklung der Sprachwahrnehmung für die weitere sprachliche Entwicklung skizziert. Der Fokus liegt auf behavioralen Befunden (für eine detaillierte Beschreibung der neurokognitiven Korrelate und Prozesse sprachlicher Entwicklung siehe Skeide & Friederici, 2016).

## Das Neugeborene

Die sprachliche Entwicklung beginnt bereits vor der Geburt (für einen Überblick zur fötalen Entwicklung siehe Einspieler et al., 2019). Neben Körpergeräuschen und der Stimme der Mutter gelangen auch Außengeräusche gefiltert in den Uterus. Gehör und Gehirn sind im letzten Schwangerschaftstrimester so weit ausgebildet, dass der Fötus diese Geräusche verarbeitet und erinnert (Michaelis, 2010). Diese pränatalen Erfahrungen prägen grundlegende Präferenzen in der Wahrnehmung kurz vor bzw.

nach der Geburt: Föten in der 38. Schwangerschaftswoche reagieren mit erhöhtem Herzschlag auf die mütterliche gegenüber einer fremden weiblichen Stimme (Kisilevsky et al., 2009) und drei Tage alte Neugeborene hören unabhängig vom Ausmaß des postnatalen Kontakts zur Mutter lieber ihre als eine fremde weibliche Stimme (DeCasper & Fifer, 1980).

Neben der Stimme werden auch spezifische Merkmale des sprachlichen Inputs verarbeitet. Neugeborene bevorzugen menschliche Sprache gegenüber komplexen nicht-sprachlichen Stimuli (Vouloumanos & Werker, 2007) und zeigen unterschiedliche neuronale Verarbeitungsmuster für normale gegenüber rückwärts gespielter Sprache (Peña et al., 2003). Neben dieser allgemeinen Präferenz für Sprache bevorzugen Neugeborene die Umgebungssprache<sup>1</sup> gegenüber einer fremden Sprache mit anderem Betonungsmuster (Moon et al., 1993). Auch diese Unterscheidungsfähigkeit von Umgebungssprache und Fremdsprache lässt sich bereits vorgeburtlich in den letzten Schwangerschaftswochen nachweisen (Kisilevsky et al., 2009). Dabei spielt die Bekanntheit der sprachlichen Betonungsmuster eine Rolle: 2-monatige englisch-lernende Säuglinge können Englisch und Italienisch, aber nicht zwei prosodisch ähnlich unterschiedliche Fremdsprachen wie Französisch und Russisch unterscheiden (Mehler et al., 1988). Gervain (2018) argumentiert daher, dass *prenatal prosodic bootstrapping* eine Grundlage der Sprachentwicklung bildet (für eine Übersicht zur weiteren Entwicklung der rhythmischen Wahrnehmung siehe Gasparini et al., 2021).

Hören und Lernen im Mutterleib erstreckt sich nicht nur auf die prosodische sondern teilweise bis auf die segmentale Ebene. Audioaufnahmen aus dem Mutterleib enthalten genügend Information, um 30 % der Laute in einem Sprachstrom zu identifizieren (Griffiths et al., 1994). Tatsächlich reagieren Föten im letzten Trimester mit Herzfrequenzmodulationen auf verschiedene Vokale (Zimmer et al., 1993) und Neugeborene zeigen innerhalb der ersten Stunden nach der Geburt ein unterschiedliches Interesse an umgebungs- gegenüber fremdsprachlichen Vokalen (Moon et al., 2013). Das zeigt, dass die pränatale Sprachexposition das Lernen phonetischer Kategorien bereits im Mutterleib anstößt und den Weg für eine sprachspezifische Wahrnehmung ebnet.

Das entstehende Wissen über die Charakteristika der Umgebungssprache hilft Neugeborenen, Informationen aus dem Sprachstrom zu extrahieren und zu verarbeiten. So können Neugeborene mehrsilbige Wörter anhand ihres Betonungsmusters unterscheiden (Sansavini et al.,

1997), verschiedene Wortklassen wie Funktions- und Inhaltswörter diskriminieren, z.B. *the* und *cat* (Shi et al., 1999); zeigen Sensitivität für akustische Merkmale, die Wortgrenzen signalisieren (Christophe et al., 1994), und für Silbenwiederholungen innerhalb von Wörtern, z.B. *mubaba* vs. *mubage* (Gervain et al., 2008). Damit ist ein wichtiger Grundstein für späteres Wortlernen und den Erwerb von grammatischen Mustern gelegt.

## Entwicklung der Sprachwahrnehmung

### Auditive Wahrnehmung

Obwohl das Neugeborene kein „unbeschriebenes Blatt“ ist, unterscheidet sich seine Sprachwahrnehmung deutlich von der eines Erwachsenen. Im Laufe des ersten Lebensjahres verändert sich die Wahrnehmung und passt sich den Besonderheiten der Umgebungssprache an. Sprachen nutzen unterschiedliche Lautinventare, z.B. kontrastiert das Englische /θ/ und /s/ wie in *thin* ‚dünn‘ vs. *sin* ‚Sünde‘. Das Deutsche enthält kein /θ/ und deutsche Muttersprachler\_innen haben ohne Training Schwierigkeiten den Unterschied akkurat zu erkennen und zu produzieren (Hanulikova & Weber, 2012). Solche sprachspezifischen Eigenheiten der Wahrnehmung und Aussprache sind Folge von Anpassungen, die im ersten Lebensjahr stattfinden: das *native language attunement* (Aslin & Pisoni, 1980).

Bereits im ersten Lebensmonat nehmen Säuglinge Sprachlaute kategorisch wahr: sie erkennen akustische Unterschiede, die eine Kategoriegrenze überschreiten, besser als Unterschiede innerhalb einer Kategorie. Eimas et al. (1971) präsentierten Säuglingen Laute auf dem Kontinuum zwischen /b/ und /p/, die sich systematisch in 20ms-Schritten in ihrer *voice onset time* (VOT) unterschieden, d.h. wann die Stimmlippen beginnen zu vibrieren. Die Säuglinge unterschieden Laute mit einem Unterschied von 20 ms in der VOT, wenn die Kategoriegrenze überschritten wurde (zB. /b/ mit 20 ms und /p/ mit 40 ms VOT), aber nicht, wenn die Laute in eine Kategorie fielen (zB. /p/ mit 60 ms und /p/ mit 80 ms VOT). Obwohl der akustische Unterschied in beiden Fällen gleich ist, reagiert das Wahrnehmungssystem unterschiedlich. Diese kategoriale Wahrnehmung ermöglicht es, Varianz im akustischen Signal zu ignorieren, die in der Umge-

<sup>1</sup> Im Artikel wird der Begriff *Umgebungssprache* anstatt *Muttersprache* verwendet. Gemeint ist damit die Sprache, mit der der Säugling in seiner Umgebung konfrontiert ist und die in der Kommunikation mit ihm verwendet wird.

bungssprache keine Bedeutung hat, während solche Unterschiede, die für die Umgebungssprache relevant sind, hervorgehoben werden.

Nachfolgende Studien zeigten, dass Säuglinge als universell Hörende geboren werden, die eine Vielzahl von Lautkontrasten der Sprachen der Welt wahrnehmen, aber für Vokale zwischen 4 und 10 Monaten (Polka & Werker, 1994; für eine Meta-Analyse siehe Tsuji & Cristia, 2014) und für Konsonanten zwischen 6 und 12 Monaten (Werker & Tees, 1999) zu sprachspezifisch Hörenden werden. So verlieren Säuglinge im Laufe des ersten Lebensjahres weitgehend die Fähigkeit fremdsprachliche Laute zu unterscheiden – mit Ausnahme solcher Laute, die sich deutlich von umgebungssprachlichen Lauten unterscheiden, wie Klicklaute einiger afrikanischer Sprachen für englische Hörende (Best et al., 1995). Auch die Wahrnehmung für umgebungssprachliche Lautkontraste verbessert sich im ersten Lebensjahr (Kuhl et al., 2006). So werden umgebungssprachlich relevante Lautkontraste anfänglich nicht sehr gut unterschieden, wenn sie akustisch wenig distinkt sind (zB. Frikative wie /s/ und /f/; Eilers et al., 1977) oder eine niedrige Auftretenswahrscheinlichkeit haben (Narayan et al., 2010). Einen aktuellen Überblick bieten hierzu Junge et al. (2019). Eine kritische Diskussion zum native language attunement und der Frage, inwieweit Säuglinge über funktionales Wissen über Lautkategorien verfügen oder Sprachlaute aufgrund phonetisch-akustischer Distanz unterscheiden, findet sich in Chladkova und Paillereau (2020).

## Audiovisuelle Wahrnehmung

Säuglinge erhalten den Großteil ihres sprachlichen Inputs in sozialer Interaktion. Der sprachliche Input ist also bimodal und enthält auditive Informationen (aus der Stimme) und visuelle Informationen (aus dem Gesicht). Der Säugling nutzt diese multimodalen Informationen, um seine Wahrnehmung an audiovisuelle Besonderheiten der Umgebungssprache anzupassen (für eine Meta-Analyse zur Wahrnehmung audiovisueller Kongruenz siehe Cox et al., 2022). Bereits mit 2 bis 4 Monaten reagieren Säuglinge auf die Kongruenz zwischen auditiven und visuellen Sprachsignalen: Sie betrachten lieber Gesichter, deren Artikulationen mit einem gehörten Laut übereinstimmen, als solche mit inkongruenten Artikulationen (Kuhl & Meltzoff, 1982). Parallel zur auditiven Sprachwahrnehmung durchläuft auch die audiovisuelle Sprachwahrnehmung im ersten Lebensjahr eine deutliche Entwicklung (Danielson et al., 2017). Während 4- bis 6-monatige Säuglinge Umgebungssprache und Fremdsprache auf Basis visueller Informationen unterscheiden können, verlieren sie im

Alter von 8 Monaten die Sensitivität für visuell-artikulatorische Unterschiede zwischen Sprachen, denen sie im Input nicht ausgesetzt sind (Weikum et al., 2007). Genauso nimmt ihre Sensitivität für die Übereinstimmung zwischen auditivem und visuellem Sprachsignal für fremdsprachliche Lautkontraste innerhalb des ersten Lebensjahres ab: 6-monatige Spanisch-lernende Säuglinge zeigen noch eine Präferenz für die mit dem gehörten Laut kongruente Artikulation des englischen Kontrasts zwischen /b/ und /v/, der im Spanischen nicht bedeutungsunterscheidend ist; 11-monatige Säuglinge und spanische Erwachsene zeigen dagegen keine Präferenz mehr (Pons et al., 2009). Gleichzeitig verfeinern Säuglinge ihre Wahrnehmung umgebungssprachlicher audiovisueller Kontraste (Altvater-Mackensen et al., 2016).

Informationen aus dem Gesicht unterstützen die auditive Sprachverarbeitung: sichtbare Mundbewegungen erleichtern Säuglingen die Wahrnehmung von Lautkontrasten (Teinonen et al., 2008) und die Segmentierung von Sprache (Hollich et al., 2005). Gemäß der *intersensory redundancy hypothesis* (Bahrick & Lickliter, 2012) erleichtert die zuverlässige Paarung von auditiven und visuellen Informationen aus derselben Quelle die Verarbeitung, indem die Aufmerksamkeit auf relevante Aspekte des Signals gelenkt wird. Tatsächlich zeigen Säuglinge in der Phase des *native language attunement* eine erhöhte Aufmerksamkeit für den Mund eines Sprechenden (Lewkowicz & Hansen-Tift, 2012), insbesondere wenn sie mehrsprachig aufwachsen (Pons et al., 2015), eine unbekannte Sprache sehen (Kubicek et al., 2013), oder visuelle und akustische Merkmale nicht kongruent sind (Tomalski et al., 2013). Diese erhöhte Aufmerksamkeit für den Mund des Sprechenden wurde mit Unterschieden in der neuronalen Sprachverarbeitung assoziiert (Altvater-Mackensen & Grossmann, 2016).

Die Sprachwahrnehmung von Säuglingen ist also von Beginn an multimodal und verknüpft auditive und visuelle Informationen (Bristow et al., 2008). Das impliziert aber nicht unbedingt, dass Säuglinge auditive und visuelle Informationen automatisch als Einheit wahrnehmen. Erwachsene integrieren konfligierende Informationen im McGurk-Effekt: ein gehörtes /b/ und gesehenes /g/ wird als /d/ wahrgenommen (McGurk & MacDonald, 1976). Diese Fähigkeit zur Integration von auditiven und visuellen Sprachstimuli ist bei Säuglingen noch nicht vollständig entwickelt (Desjardins & Werker, 2004). Eine kritische Diskussion der sich entwickelnden Fähigkeit zur audiovisuellen Integration liefern Shaw und Bortfeld (2015).

## Einflussfaktoren auf die Entwicklung der Sprachwahrnehmung

### Der Zusammenhang von Sprachproduktion und -perzeption

Parallel zur Entwicklung der Sprachwahrnehmung beginnen auch die Vokalisationen umgebungssprachliche Merkmale zu zeigen (einen systematischen Überblick über die Entwicklung frühkindlicher Produktionen bieten Morgan & Wren, 2018). Einige Forscher postulieren daher eine enge Verknüpfung der Entwicklung über Domänen hinweg. So nimmt Vihman (1996) in der *Articulatory Filter Hypothesis* an, dass die frühkindliche Sprachwahrnehmung durch artikulatorische Fähigkeiten moduliert wird. Durch einen automatisch stattfindenden Abgleich mit motorischen Mustern treten Lautsequenzen, die der Säugling bereits selbst produziert, in der Wahrnehmung besonders salient hervor. Evidenz für eine solche Beeinflussung der Wahrnehmung durch motorische Fähigkeiten finden Majorano et al. (2014): Säuglinge zeigen zu Beginn des Babbels in der Wahrnehmung eine Präferenz für Lautsequenzen, die sie selbst produzieren; später verschiebt sich ihre Präferenz zu Lautmustern, die außerhalb ihres Repertoires liegen. Säuglinge mit größerem produktiven Lautinventar zeigen außerdem eine höhere Sensitivität für die Kongruenz audiovisueller Sprache (Altvater-Mackensen et al., 2016). Dabei werden individuelle Unterschiede in der audiovisuellen Wahrnehmung nicht durch die Auftretenswahrscheinlichkeit der einzelnen Laute in der Umgebungssprache bestimmt, sondern durch die produktiven Fähigkeiten des Säuglings (Streri et al., 2016). Einen unmittelbaren Einfluss motorischen Feedbacks auf die Wahrnehmung zeigen Yeung und Werker (2013): Sie präsentierten 4- bis 5-monatigen Säuglingen zwei Gesichter, die stumm /i/ und /u/ artikulierten; der gehörte Laut war kongruent mit einer der Artikulationen. Gleichzeitig lutschten die Säuglinge an einem Schnuller oder Ring, um motorisches Feedback zu induzieren, das der Produktion von /o/ (Schnuller) oder /i/ (Ring) entspricht. Die Ergebnisse zeigen, dass die audiovisuelle Verarbeitung durch die konkurrierende motorische Information beeinflusst wird, und legen damit eine Verknüpfung von auditorischer, visueller und artikulatorischer Information in der Sprachwahrnehmung nahe (für eine kritische Diskussion sensomotorischen Feedbacks in der Sprachwahrnehmung siehe Guellai et al., 2014).

### Quantität und Qualität des Inputs

Der Input spielt eine wichtige Rolle für die Entwicklung der Sprachwahrnehmung. Gut untersucht ist *infant-directed speech* (IDS), ein an die Kommunikation mit Säuglingen angepasster Sprechstil in hoher Tonlage mit starker Intonation, langsamer und deutlicher Artikulation und kurzen Sätzen (Fernald et al., 1989). Die spezifischen Eigenheiten von IDS können die Sprachverarbeitung erleichtern (Zangl & Mills, 2007) und die Aufmerksamkeit des Säuglings auf für ihn relevanten Sprachinput richten (Cooper & Aslin, 1990). Säuglinge, deren Mütter ausgeprägtes IDS in der Kommunikation nutzen, zeigen eine höhere Sensitivität für umgebungssprachliche Lautkontraste als Säuglinge, deren Mütter ihre Sprache nur wenig anpassen (Cristia, 2013). Forschende haben daher argumentiert, dass phonologische Kontraste der Umgebungssprache durch IDS verstärkt werden und das *perceptual attunement* erleichtern (Liu et al., 2003). Da IDS Lautkontraste aber verzerrt, könnte der besondere Sprechstil auch primär der Fokussierung der Aufmerksamkeit (McMurray et al., 2013) oder der Vermittlung von positivem Affekt dienen (Benders, 2013). Eine detaillierte Beschreibung über Eigenschaften und Wahrnehmung von IDS sowie ihre Bedeutung für die sprachliche Entwicklung liefert Cristia (2013).

Neben IDS ist die reziproke Kommunikation in sozialen Interaktionen wichtig für die frühe Sprachentwicklung (Golinkoff et al., 2018). Der Säugling ist kein passiver Zuhörer, sondern sendet selbst kommunikative Signale aus, die von den Bezugspersonen erkannt und intuitiv beantwortet werden (Papoušek, 2007). Die kontingente und angemessene Reaktion auf das Verhalten des Säuglings ist für das Lernen wesentlich. Säuglinge mit kontingent reagierenden Müttern zeigen ein früheres *perceptual attunement* (Elsabbagh et al., 2013), eine höhere Sensitivität für audiovisuelle Kongruenz (Altvater-Mackensen & Grossmann, 2015) und mehr zielsprachliche Vokalisationen (Goldstein et al., 2003). Kuhl et al. (2003) belegen die zentrale Rolle sozialer Interaktion: 9-monatige Englischlernende Säuglinge erhielten mehrwöchigen Input von einer Mandarin-Sprecherin. Eine Gruppe interagierte direkt mit der Sprecherin, eine zweite sah Videoaufnahmen, eine dritte hörte Tonaufzeichnungen. Obwohl der akustische Sprachinput für alle Gruppen identisch war, diskriminierten nur die Säuglinge in der Interaktionsgruppe später Lautkontraste des Mandarin. Qualitative Merkmale des Inputs wirken also über quantitative Merkmale hinaus auf die sprachliche Entwicklung.

Auswirkungen von Unterschieden im Input zeigen sich auch in der Entwicklung mehrsprachig aufwachsender Säuglinge. Obwohl sie weniger Input in jeder einzelnen Sprache erhalten, nehmen sie Kontraste ihrer Umgebungssprachen am Ende des ersten Lebensjahres ähnlich



gut wahr wie einsprachig aufwachsende Säuglinge (Sundara et al., 2008). Das *perceptual attunement* kann jedoch eine U-förmige Entwicklung mit zeitweise verringerter Sensitivität zeigen (Sebastian-Galles, 2010) und nur die Verarbeitung der dominanten Sprache zeigt vergleichbare neuronale Muster zu einsprachigen Säuglingen (Garcia-Sierra et al., 2011). Außerdem zeigen mehrsprachige Säuglinge länger Sensitivität für fremdsprachliche Lautkontraste sowohl in der auditorischen (Singh et al., 2017), als auch in der visuellen Sprachwahrnehmung (Sebastian-Galles et al., 2012). Die *Perceptual Wedge Hypothesis* (Pettito et al. 2012) nimmt daher an, dass mehrsprachiger Input zu einer erhöhten Flexibilität der Sprachwahrnehmung führt. Wie Maurer und Werker (2014) diskutieren, spricht eine erhöhte Sensitivität für fremdsprachliche Lautkontraste allerdings nicht notwendigerweise für Unterschiede im Entwicklungsverlauf des *perceptual attunement*, sondern könnte auch auf Unterschieden in Aufmerksamkeits- und Verarbeitungsprozessen beruhen, z. B. verbesserte exekutive Funktionen (Kovacs & Mehler, 2009). So finden Liu und Kager (2016) eine erhöhte Sensitivität für umgebungssprachliche Kontraste in mehrsprachig gegenüber einsprachig aufwachsenden Säuglingen und argumentieren, dass variabler Input zu einer generell erhöhten Sensitivität für systematische phonetische Unterschiede führt. Einen aktuellen Überblick über Variabilität und Stabilität in der Verarbeitung phonologischer Information durch ein- und mehrsprachig aufwachsende Säuglinge geben Höhle et al. (2020).

## Modelle frühkindlicher Sprachwahrnehmung

Wie lernen Säuglinge welche Lautkontraste in der Umgebungssprache relevant sind und welche Unterschiede ignoriert werden können? Aslin und Pisoni (1980) nehmen in der *attunement theory* an, dass sich die Grundlagen der Sprachwahrnehmung pränatal entwickeln. Nach der Geburt gibt es drei Entwicklungsmuster: 1) einen Verlust der Fähigkeit, Sprachlaute zu diskriminieren, die in der Umgebungssprache fehlen, z. B. /u/ und /y/ in *Mutter* vs. *Mütter* für Englisch-Lernende (Polka & Werker, 1994); 2) eine Verbesserung der Wahrnehmungsfähigkeit für Laute, die zu Beginn nur schlecht wahrgenommen werden, aber in der Umgebungssprache relevant sind, wie der Kontrast zwischen einzelnen Frikativen (Eilers et al., 1977); 3) ein Erhalt der Wahrnehmungsfähigkeit für Lautkontraste, die von Beginn an gut wahrgenommen werden und zum Inventar der Umgebungssprache gehören. Wie postnatale Erfahrungen die Entwicklung formen, ist eine

offene theoretische Debatte (siehe Werker & Hensch, 2015, für eine Diskussion sensitiver Phasen und ihrer biologischen Grundlagen).

## Statistisches Lernen und das sich entwickelnde Lexikon

Eine weithin akzeptierte Antwort, wie der Input Veränderungen in der Wahrnehmung beeinflusst, ist statistisches Lernen. Jede Realisierung eines Phonems unterscheidet sich (z. B. aufgrund von Koartikulation; Peterson & Barney, 1952). Die resultierende Variation ist nicht zufällig, sondern zeigt lokale Maxima entlang relevanter phonetischer Dimensionen, die Sprachlaute voneinander unterscheiden (Allen & Miller, 1999). Solche statistischen Verteilungen modulieren die Wahrnehmung von Sprachlauten (Maye et al., 2002). Die Idee des statistischen Lernens als Mechanismus, der das *perceptual attunement* vorantreibt, findet sich daher in vielen Theorien der frühen Sprachwahrnehmung (z. B. Werker & Curtin, 2005; Kuhl et al., 2008).

Simulationen des Entwicklungsprozesses in Modellen legen jedoch nahe, dass Lautkategorien nur unter Berücksichtigung des Wortkontexts erfolgreich gelernt werden können (Feldman et al., 2013). Verhaltensstudien stützen diese Annahme: Säuglingen fällt es leichter Sprachlaute zu unterscheiden, wenn sie zuvor in verschiedenen Wortkontexten präsentiert wurden (z. B. /p/ in Puppe und /t/ in Teddy; Thiessen, 2011). Generell verbessert die konsistente Paarung von Sprachlauten mit unterschiedlichen Stimuli wie Gesichtern (Teinonen et al., 2008) oder Objekten (Yeung & Werker, 2009) die Diskriminationsfähigkeit im Sinne einer *acquired distinctiveness* (Hall, 1991). Da Säuglinge mit 6 bis 9 Monaten erste Wörter verstehen (Tincoff & Juszyck, 2012), könnte das sich entwickelnde Lexikon *perceptual attunement* unterstützen. Laut Swingley (2009) hilft das Wissen über Wortformen im Proto-Lexikon Säuglingen, mit der Varianz im auditorischen Input umzugehen: sie gewichten bedeutungsunterscheidende Lautkontraste (/u/ und /o/ in *Mund* vs. *Mond*) stärker als Varianz, die nicht bedeutungsunterscheidend ist (unterschiedliche Realisationen von /u/ in *Mund*), und etablieren so umgebungssprachliche Lautkategorien. Thiessen und Pavlik (2016) zeigen anhand des typischen Wortschatzes englischer Kleinkinder, dass der Input so strukturiert ist, dass das kontrastive Erlernen von Lautkategorien im Prinzip möglich ist. Dennoch findet eine Meta-Analyse von Bergmann et al. (2017) keine Evidenz dafür, dass ein Proto-Lexikon die Grundlage für den Erwerb umgebungssprachlicher Lautkategorien bildet.

Eine daran anschließende Frage ist, inwieweit sich lautsprachliche Repräsentationen grundlegend durch den Erwerb eines Lexikons verändern (siehe *lexical restructuring model*, Metsala & Walley, 1998). Das PRIMIR-Modell (Werker & Curtin, 2005) geht davon aus, dass sich die lautliche Information verändert, die im Laufe der sprachlichen Entwicklung genutzt wird. Diese Annahme basiert auf scheinbar widersprüchlichen Befunden, dass Säuglinge zwar bereits im ersten Lebensjahr verschiedenste Lautkontraste unterscheiden können, im zweiten Lebensjahr aber Schwierigkeiten haben, dieses lautliche Detail beim Wortlernen zu nutzen (Stager & Werker, 1997). PRIMIR nimmt unterschiedliche multidimensionale und interaktive Repräsentationsebenen an, die im Zusammenspiel mit den kognitiven Anforderungen die Wahrnehmung filtern. Zu den Repräsentationsebenen gehört eine allgemein-perzeptuelle Ebene, eine Wortform- und eine Phonem-Ebene. Informationen des Inputs, wie phonetisch-akustische Information, Index-Information (z.B. Geschlecht eines Sprechenden) oder visuelle Information werden in der allgemein-perzeptuellen Ebene gespeichert. Exemplar-basierte Wortformen (d.h. in Sequenzen gehäuft zusammen auftretende Laute) werden auf der Wortform-Ebene gespeichert. Gemeinsame Auftretenswahrscheinlichkeiten ergeben Gruppierungen und Cluster, die durch statistisches Lernen die Assoziation von Wortformen und Bedeutungen erlauben. Ähneln sich die Lautformen von Wörtern, stehen dem Kind nicht ausreichend kognitive Ressourcen zur Verfügung, um neben der Form-Bedeutungs-Relation auch phonetisch spezifizierte Formen zu repräsentieren. Als Konsequenz wird ein Lautkontrast, der zwar grundsätzlich diskriminiert werden kann, nicht direkt im lexikalischen Lernen genutzt. Die Phonem-Ebene repräsentiert phonologische Kategorien, die auf Basis des sich entwickelnden Lexikons abgeleitet werden. Laut diesem Modell basiert die Sprachwahrnehmung zu Beginn auf phonetisch-akustischen Informationen, während phonemische Kategorien, also linguistisch bedeutungstragende Einheiten, sich erst im Laufe der Entwicklung ausbilden (für einen statistischen Ansatz siehe Feldmann et al., 2021).

Zusammengenommen ist es umstritten, wie der Erwerb von umgebungssprachlichen Lautkategorien und die lexikalische Entwicklung sich gegenseitig beeinflussen und voneinander abhängen.

## Perzeptuelles Lernen und multimodale Informationen

Ein multimodales Modell des *perceptual attunement* schlagen Westermann und Miranda (2004) vor. Es konzentriert sich auf die Rolle des Babbels und die Ver-

knüpfung motorischer und sensorischer Lautinformationen. In ihrem konnektionistischen Modell werden Sprachlaute in zwei sich selbst organisierenden *maps* repräsentiert, die einer auditorischen und einer motorischen Ebene entsprechen. Verbindungen zwischen den Domänen entstehen durch gleichzeitige Aktivierung in beiden Ebenen. Hört ein Säugling Sprachlaute, werden die entsprechenden Repräsentationen der auditorischen Ebene aktiviert, die z.B. für Vokale Formanten und ihre Frequenzen abbilden. Produziert der Säugling selbst Sprachlaute, werden entsprechende Repräsentationen in der motorischen Ebene aktiviert, die Artikulatoren und ihre Position abbilden. Der resultierende selbst produzierte Laut aktiviert wiederum Repräsentationen der auditorischen Ebene. So dienen die eigenen Produktionen dem Säugling als Feedback, um multimodale Verbindungen zu schaffen. So bilden sich auf Basis multimodalen statistischen Lernens umgebungssprachliche Lautkategorien aus.

Eine perzeptions-orientierte Sichtweise nehmen Kuhl et al. (2008) in der *Native Language Magnet Theory* ein, laut der es drei Entwicklungsphasen gibt: 1) Zu Beginn unterscheiden Säuglinge Sprachlaute aufgrund allgemeiner auditiver Fähigkeiten. 2) Durch statistisches Lernen bilden sich Kategorien heraus, die die Wahrnehmung formen. Säuglinge werden weniger sensitiv für Unterschiede innerhalb von Kategorien und sensitiver für Unterschiede an Kategoriengrenzen. Mit zunehmender Erfahrung funktionieren die prototypischen Exemplare einer Kategorie wie Magneten, die andere Kategorienmitglieder anziehen und so innerkategoriale Unterschiede weiter verringern. 3) Das resultiert in erhöhter Sensitivität für umgebungssprachliche und veringert Sensitivität für fremdsprachliche Kontraste (*perceptual magnet effect*). Die Veränderungen in der Sprachwahrnehmung sind das Resultat neuronaler Bahnung, die aufgrund wiederkehrender Muster im Input entsteht und das weitere Lernen beeinflusst. Hierbei spielen soziale Interaktion und IDS eine besondere Rolle, weil sie umgebungssprachliche Lautkontraste hervorheben. Entsprechend der *social gating hypothesis* (Kuhl, 2007) helfen sie, die kindliche Aufmerksamkeit zu fokussieren und den Säugling lernbereit zu machen. Die frühe Entwicklung ist demnach also primär durch Perzeptionsprozesse beeinflusst. Der Link zwischen Perzeption und Produktion wird erst im Laufe der Entwicklung durch Imitationsprozesse und die Verarbeitung des eigenen Inputs geformt.

Eine motorisch-orientierte Sichtweise nehmen im Kontrast dazu Best und McRoberts (2003) im *perceptual assimilation model* und der *articulatory organ matching hypothesis* ein, nach denen Lautkategorien als motorische Gesten repräsentiert und interpretiert werden. Perzeption und Produktion greifen auf universelle motorische Kate-

gorien zurück, die keine Interpretation des akustischen Inputs sind, sondern direkt als solche identifiziert werden. Die linguistisch-funktionale Bedeutung dieser Kategorien bildet sich erst während des ersten Lebensjahres auf Basis des sprachlichen Inputs heraus. Die Wahrnehmung fremdsprachlicher Laute hängt von ihrem Verhältnis zu Lautkategorien der Umgebungssprache ab (Best et al., 2016). Überlappen die Laute mit umgebungssprachlichen Artikulationskategorien oder sind ihnen hinreichend ähnlich, werden sie assimiliert. Weichen sie stark von umgebungssprachlichen Kategorien ab, bleiben sie unkategorisiert. Sind zwei fremdsprachliche Laute mit den gleichen Artikulatoren assoziiert, assimilieren sie an die gleiche umgebungssprachliche Kategorie und werden nicht mehr unterschieden. Fremdsprachliche Laute können dagegen weiter unterschieden werden, wenn sie an verschiedene Kategorien assimilieren oder unkategorisiert bleiben. So erklärt das Modell unterschiedliche Sensitivitäten für fremdsprachliche Laute nach dem *perceptual attunement* entsprechend ihrer artikulatorischen Eigenschaften und nimmt einen inhärenten Zusammenhang zwischen Produktion und Perzeption an.

Zusammenfassend postulieren die vorgestellten Modelle verschiedene Mechanismen, durch die sich umgebungssprachliche Lautkategorien entwickeln und so die Sprachwahrnehmung im Laufe der Entwicklung verändern. Die Modelle unterscheiden sich dahingehend, ob sie Produktionsprozessen dabei ein inhärente Rolle zuschreiben oder nicht, und ob sie einen qualitativen Wechsel von phonetisch-akustischer zu phonemischer Wahrnehmung annehmen. Existierende Modelle fokussieren sich dabei weitgehend auf die auditive Domäne, Modelle der frühen sprachlichen Entwicklung, die multimodale Informationen auch über artikulatorische und auditive Information hinaus einbinden, fehlen weitgehend.

## Bedeutung früher Sprachwahrnehmung für die weitere Sprachentwicklung

Auch wenn bislang keine Einigkeit über die Mechanismen des *perceptual attunement* besteht, so ist seine Bedeutung für die weitere Sprachentwicklung unumstritten. Durch die Veränderungen wird der Säugling zu einem *native listener*, dessen Perzeptionsprozesse für die Verarbeitung der Umgebungssprache optimiert sind. So zeigen sich Korrelationen zwischen fortgeschrittener perzeptueller Entwicklung im ersten Lebensjahr und Größe des Wortschatzes im zweiten Lebensjahr für die auditorische (Tsao et al., 2004) und die audiovisuelle Sprachwahrnehmung

(Altvater-Mackensen & Grossmann, 2015). Dabei scheint der Erwerb von Konsonanten- aber nicht von Vokalkategorien prädiktiv für die weitere lexikalische Entwicklung zu sein (Schaadt et al., 2023). Ein Befund, der im Einklang steht mit der *consonant vowel hypothesis* (Nespor et al., 2003), nach der Konsonanten informativer für die lexikalische Verarbeitung sind als Vokale.

Weiterhin scheint es einen Zusammenhang zwischen Lautwahrnehmung und schriftsprachlicher Entwicklung zu geben. Der Erwerb von Lese- und Schreibfähigkeiten erfordert in alphabetischen Schriftsystemen die Assoziation von Phonemen und Graphemen. Eine Grundvoraussetzung dafür ist die effiziente Diskrimination und Kategorisierung von Sprachlauten, die bei Dyslexie beeinträchtigt ist (de Gelder & Vroomen, 1998). Allerdings zeigen sich Beeinträchtigungen in der Wahrnehmung von sprachlichen und nicht-sprachlichen Lautkontrasten (Vandermosten et al., 2010), so dass ein gemeinsames Defizit in der allgemein-auditorischen Verarbeitung angenommen wird (Tallal, 1980). Doch sind auch bei typischer Entwicklung kategoriale Wahrnehmung und Identifikation von Lauten bei Hintergrundrauschen zu Beginn des Schriftspracherwerbs Prädiktoren für die weitere Lese- und Schreibfähigkeit (Vandewalle et al., 2012).

Die Vorteile einer an die Umgebungssprache angepassten Lautwahrnehmung fordern ihren Preis in der Schwierigkeit, fremdsprachliche Lautkontraste akkurat wahrzunehmen und zu produzieren. Frühe Erfahrungen beeinflussen langfristig selbst dann noch die Wahrnehmungsfähigkeit, wenn der Säugling einer Sprache nur zeitweise ausgesetzt war: Niederländische Erwachsene, die mit wenigen Monaten aus Korea adoptiert wurden, konnten koreanische Lautkontraste nach kurzem Training besser unterscheiden als eine Kontrollgruppe, die in den Niederlanden ohne koreanischen Input aufwuchs (Choi et al., 2017). Dennoch ist das Erlernen von fremdsprachlichen Lautkategorien bei entsprechendem Training auch im Erwachsenenalter noch möglich (Bradlow et al., 1999).

## Fazit

Bereits bevor sie ihr erstes Wort sprechen, erreichen Säuglinge wichtige Meilensteine der sprachlichen Entwicklung. Im Laufe des ersten Lebensjahres verändert sich ihre Sprachwahrnehmung und stellt sich auf die Umgebungssprache ein, zunächst auf die prosodischen dann auf die segmentalen Eigenschaften. Welche Mechanismen und Prozesse dieses *perceptual attunement* formen, ist noch nicht vollständig geklärt. Sowohl quantitative als auch qualitative Merkmale des sprachlichen In-

puts und der sozialen Interaktion mit dem Säugling, als auch die Entwicklung artikulatorischer Fähigkeiten beeinflussen die frühkindliche Sprachwahrnehmung. Modelle und Theorien des *perceptual attunement* legen hier unterschiedliche Schwerpunkte und benennen sowohl *bottom-up* Prozesse, wie die Verarbeitung statistischer Informationen oder multimodaler Verknüpfungen, als auch *top-down* Prozesse, wie kontrastive Information im sich entwickelnden Lexikon, als treibende Faktoren in der Entwicklung. Zusammengenommen zeigt das Wahrnehmungssystem eine hohe Anpassungsfähigkeit, die eine mühelose Sprachverarbeitung erlaubt und eine Grundlage für die weitere sprachliche Entwicklung bildet.

## Literatur

- Allen, J. S. & Miller, J. L. (1999). Effects of syllable-initial voicing and speaking rate on the temporal characteristics of monosyllabic words. *Journal of the Acoustical Society of America*, 106, 2031–2039. <https://doi.org/10.1121/1.427949>
- Altvater-Mackensen, N. & Grossmann, T. (2015). Learning to match auditory and visual speech cues: Social influences on the acquisition of phonological categories. *Child Development*, 86, 362–378. <https://doi.org/10.1111/cdev.12320>
- Altvater-Mackensen, N. & Grossmann, T. (2016). The role of left inferior frontal cortex during audiovisual speech perception in infants. *NeuroImage*, 133, 14–20. <https://doi.org/10.1016/j.neuroimage.2016.02.061>
- Altvater-Mackensen, N., Mani, N. & Grossmann, T. (2016). Audiovisual speech perception in infancy: The influence of vowel identity and productive abilities on infants' sensitivity to (mis) matches between auditory and visual speech cues. *Developmental Psychology*, 52, 191–204. <https://doi.org/10.1037/a0039964>
- Aslin, R. N. & Pisoni, D. B. (1980). Some developmental processes in speech perception. In G. H. Yeni-Komshian, J. F. Kavanaugh & C. A. Ferguson (Eds.), *Child phonology* (Vol. 2, pp.67–96). New York, NY: Academic.
- Bahrack, L. E. & Lickliter, R. (2012). The role of intersensory redundancy in early perceptual, cognitive and social development. In J. Bremner, D.J. Lewkowicz & C. Spence (Eds.), *Multisensory development* (pp. 183–205). Oxford, UK: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199586059.003.0008>
- Benders, T. (2013). Mommy is only happy! Dutch mothers' realization of speech sounds in infant-directed speech expresses emotion, not didactic intent. *Infant Behavior and Development*, 36(4), 847–862. <https://doi.org/10.1016/j.infbeh.2013.09.001>
- Bergmann, C., Tsuji, S. & Cristia, A. (2017). Top-down versus bottom-up theories of phonological acquisition: A big data approach. *Proceedings of Interspeech 2017*, 2103–2107. <https://doi.org/10.21437/Interspeech.2017-1443>
- Best, C. T., McRoberts, G. W., LaFleur, R. & Silver-Isenstadt, J. (1995). Divergent developmental patterns for infants' perception of two nonnative consonant contrasts. *Infant Behavior and Development*, 18, 339–350. [https://doi.org/10.1016/0163-6383\(95\)90022-5](https://doi.org/10.1016/0163-6383(95)90022-5)
- Best, C. T. & McRoberts, G.W. (2003). Infant perception of nonnative consonant contrasts that adults assimilate in different ways. *Language and Speech*, 46(2-3), 183–216. <https://doi.org/10.1177/00238309030460020701>
- Best, C. T., Goldstein, L. M., Nam, H. & Tyler, M. D. (2016). Articulating what infants attune to in native speech. *Ecological Psychology*, 28(4), 216–261. <https://doi.org/10.1080/10407413.2016.1230372>
- Bradlow, A., Akahane-Yamada, R., Pisoni, D. B. & Tohkura, Y. (1999). Training Japanese listeners to identify English /r/ and /l/: Long-term retention of learning in perception and production. *Perception & Psychophysics*, 61(5), 977–985. <https://doi.org/10.3758/BF03206911>
- Bristow, D., Dehaene-Lambertz, G., Mattout, J., Soares, C., Gliga, T., Baillet, S. & Mangin, J. F. (2008). Hearing faces: How the infant brain matches the face it sees with the speech it hears. *Journal of Cognitive Neuroscience*, 21, 905–921. <https://doi.org/10.1162/jocn.2009.21076>
- Chladkova, K. & Paillereau, N. (2020). The what and when of universal perception: A review of early speech sound acquisition. *Language Learning*, 70, 1136–1182. <https://doi.org/10.1111/lang.12422>
- Choi, J., Broersma, M. & Culter, A. (2017). Early phonology revealed by international adoptees' birth language retention. *Proceedings of the National Academy of Sciences*, 114(28), 7307–7312. <https://doi.org/10.1073/pnas.1706405114>
- 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. <https://doi.org/10.1121/1.408544>
- Cooper, R.P. & Aslin, R.N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development*, 61, 1584–1595. <https://doi.org/10.1111/j.1467-8624.1990.tb02885.x>
- Cox, M. M., Keren-Portnoy, T., Roepstorff, A. & Fusaroli, R. (2022). A Bayesian meta-analysis of infants' ability to perceive audiovisual congruence for speech. *Infancy*, 27, 67–96. <https://doi.org/10.1111/infa.12436>
- Cristia, A. (2013). Input to language: The phonetics and perception of infant-directed speech. *Language and Linguistics Compass*, 7, 157–170. <https://doi.org/10.1111/lnc3.12015>
- Danielson, K., Bruderer, A. G., Kandhadai, P., Vatikiotis-Bateson, E. & Werker, J. F. (2017). The organization and reorganization of audiovisual speech perception in the first year of life. *Cognitive Development*, 42, 37–48. <https://doi.org/10.1016/j.cogdev.2017.02.004>
- DeCasper, A. J. & Fifer, W. P. (1980). Of human bonding. Newborns prefer their mothers' voices. *Science*, 208, 1174–1176. <https://doi.org/10.1126/science.7375928>
- de Gelder, B. & Vroomen, J. (1998). Impaired speech perception in poor readers: Evidence from hearing and speech reading. *Brain and Language*, 64, 269–281. <https://doi.org/10.1006/brln.1998.1973>
- Desjardins, R. N. & Werker, J. F. (2004). Is the integration of heard and seen speech mandatory for infants? *Developmental Psychobiology*, 45, 187–203. <https://doi.org/10.1002/dev.20033>
- Eimas, P. D., Siqueland, E. R., Jusczyk, P. W. & Vigorito, J. (1971). Speech perception in infants. *Science*, 171(968), 303–306. <https://doi.org/10.1126/science.171.3968.3>
- Eilers, R. E., Wilson, W. R. & Moore, J. M. (1977). Developmental changes in speech discrimination in infants. *Journal of Speech and Hearing Research*, 20, 766–780. <https://doi.org/10.1044/jshr.2004.766>
- Einspieler, C., Prayer, D. & Prechtel, H. (2019). *Fetal behavior: A neurodevelopmental approach*. London, UK: MacKeith.
- Elsabbagh, M., Hohenberger, A., Campos, R., Van Herwegen, J., Serres, J., de Schonen, S., Aschersleben, G. & Karmiloff-Smith, A. (2013). Narrowing perceptual sensitivity to the native language in infancy: Exogenous influences on developmental ti-



- ming. *Behavioural Sciences*, 3, 120–132. <https://doi.org/10.3390/bs3010120>
- Feldman, N. H., Goldwater, S., Griffiths, T. & Morgan, J. L. (2013). A role for the developing lexicon in phonetic category acquisition. *Psychological Review*, 120, 751–778. <https://doi.org/10.1037/a0034245>
- Feldman, N. H., Goldwater, S., Dupoux, E. & Schatz, T. (2021). Do infants really learn phonetic categories? *Open Mind*, 5, 113–131. [https://doi.org/10.1162/opmi\\_a\\_00046](https://doi.org/10.1162/opmi_a_00046)
- Fernald, A., Taeschner, T., Dunn, J., Papoušek, M., de Boysson-Bardies, B. & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language*, 16, 477–501. <https://doi.org/10.1017/S0305000900010679>
- García-Sierra, A., Rivera-Gaxiola, M., Percaccio, C. R., Conboy, B. T., Romo, H., Klarman, L., Ortiz, S. & Kuhl, P. K. (2011). Bilingual language learning: An ERP study relating early brain responses to speech, language input, and later word production. *Journal of Phonetics*, 39, 546–557. <https://doi.org/10.1016/j.wocn.2011.07.002>
- Gasparini, L., Langus, A., Tsuji, S. & Boll-Avetisyan, N. (2021). Quantifying the role of rhythm in infants' language discrimination abilities: A meta-analysis. *Cognition*, 213, 104757. <https://doi.org/10.1016/j.cognition.2021.104757>
- Gervain, J. (2018). The role of prenatal experience in language development. *Current Opinions in Behavioral Sciences*, 21, 62–67. <https://doi.org/10.1016/j.cobeha.2018.02.004>
- Gervain, J., Macagno, F., Cogoi, S., Pena, M. & Mehler, J. (2008). The neonate brain detects speech structure. *Proceedings of the National Academy of Sciences*, 105, 14222–14227. <https://doi.org/10.1073/pnas.0806530105>
- Goldstein, M. H., King, A. P. & West, M. J. (2003). Social interaction shapes babbling: Testing parallels between birdsong and speech. *Proceedings of the National Academy of Sciences*, 100, 8030–8035. <https://doi.org/10.1073/pnas.1332441100>
- Golinkoff, R. M., Hoff, E., Rowe, M. L., Tamis-LeMonda, C. S. & Hirsh-Pasek, K. (2018). Language matters: Denying the existence of the 30-million-word-gap has serious consequences. *Child Development*, 90, 985–992. <https://doi.org/10.1111/cdev.13128>
- Griffiths, S. K., Brown, W. S., Gerhardt, K. J., Abrams, R. M. & Morris, R. (1994). The perception of speech sound recorded within the uterus of a pregnant sheep. *Journal of the Acoustical Society of America*, 96, 2055–2063. <https://doi.org/10.1121/1.410147>
- Guellai, B., Streri, A. & Yeung, H. H. (2014). The development of sensorimotor influences in the audiovisual speech domain: Some critical questions. *Frontiers in Psychology*, 5, 812. <https://doi.org/10.3389/fpsyg.2014.00812>
- Hall, G. (1991). *Perceptual and associative learning*. Oxford: Oxford University Press.
- Hanulíková, A. & Weber, A. (2012). Sink positive: Linguistic experience with the substitutions influences nonnative word recognition. *Attention, Perception & Psychophysics*, 74, 613–629. <https://doi.org/10.3758/s13414-011-0259-7>
- Hellrung, U. (2019). *Sprachentwicklung und Sprachförderung. Beobachten, verstehen, handeln*. Freiburg: Herder.
- Höhle, B., Bijeljic-Babic, R. & Nazzi, T. (2020). Variability and stability in early language acquisition: Comparing monolingual and bilingual infants' speech perception and word recognition. *Bilingualism: Language and Cognition*, 23, 56–71. <https://doi.org/10.1017/S1366728919000348>
- Hollich, G., Newman, R. S. & Juszyk, P. W. (2005). Infants' use of synchronized visual information to separate streams of speech. *Child Development*, 76, 598–613. <https://doi.org/10.1111/j.1467-8624.2005.00866.x>
- Junge, C., Boll-Avetisyan, N. & Benders, T. (2019). Speech perception and discrimination: From sounds to words. In J. S. Horst & J. von Koss Torkildsen (Hrsg.), *International handbook of language acquisition* (pp. 153–172). New York, NY: Routledge.
- Kauschke, C. (2012). *Kindlicher Spracherwerb im Deutschen: Verläufe, Forschungsmethoden, Erklärungsansätze*. Berlin: De Gruyter.
- Kisilevsky, B. S., Hains, S. M. J., Brown, C. A., Lee, C. T., Cowperthwaite, B., Stutzman, S. S., Swansburg, M. L., Lee, K., Xie, X., Huang, H., Ye, H.-H., Zhand, W. & Wang, Z. (2009). Fetal sensitivity to properties of maternal speech and language. *Infant Behaviour and Development* 32(1), 59–71. <https://doi.org/10.1016/j.infbeh.2008.10.002>
- Kovacs, A. M. & Mehler, J. (2009). Cognitive gains in 7-month-old bilingual infants. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 6556–6560. <https://doi.org/10.1073/pnas.0811323106>
- Kubicek, C., Hillairet de Boisferon, A. & Schwarzer, G. (2013). Face-scanning behavior to silently-talking faces in 12-month-old infants: The impact of pre-exposed auditory speech. *International Journal of Behavioral Development*, 37(2), 106–110. <https://doi.org/10.1177/0165025412473016>
- Kuhl, P. K. (2007). Is speech learning 'gated' by the social brain? *Developmental Science*, 10, 110–120. <https://doi.org/10.1111/j.1467-7687.2007.00572.x>
- Kuhl, P. K. & Meltzoff, A. N., (1982). The bimodal perception of speech in infancy. *Science*, 218, 1138–1141. <https://doi.org/10.1126/science.7146899>
- Kuhl, P. K., Tsao, F. & Liu, H. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences*, 100, 9096–9101. <https://doi.org/10.1073/pnas.1532872100>
- Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S. & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science*, 9, F13–21. <https://doi.org/10.1111/j.1467-7687.2006.00468.x>
- Kuhl, P. K., Conboy, B., Coffey-Corina, S., Padden, D., Riera-Gaxiola, M. & Nelson, T. (2008). Phonetic learning as a pathway to language: New data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society B*, 363, 979–1000. <https://doi.org/10.1098/rstb.2007.2154>
- Lewkowicz, D. J. & Hansen-Tift, A. M. (2012). Infants deploy selective attention to the mouth of a talking face when learning speech. *Proceedings of the National Academy of Sciences*, 109, 1431–1436. <https://doi.org/10.1073/pnas.1114783109>
- Liu, H., Kuhl, P. K. & Tsao, F. (2003). An association between mothers' speech clarity and infants' speech discrimination skills. *Developmental Science*, 6, F1-F10. <https://doi.org/10.1111/1467-7687.00275>
- Liu, L. & Kager, R. (2016). Perception of a native vowel contrast by Dutch monolingual and bilingual infants: A bilingual perceptual lead. *International Journal of Bilingualism*, 20, 335–345. <https://doi.org/10.1177/1367006914566082>
- Majorano, M., Vihman, M. M. & DePaolis, R. A. (2014). The relationship between infants' production experience and their processing of speech. *Language Learning and Development*, 10, 179–204. <https://doi.org/10.1080/15475441.2013.829740>
- Maurer, D. & Werker, J. F. (2014). Perceptual Narrowing during infancy: A comparison of language and faces. *Developmental Psychobiology*, 56(2), 154–178. <https://doi.org/10.1002/dev.21177>
- Maye, J., Werker, J. F. & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, 82, B101–B111. [https://doi.org/10.1016/s0010-0277\(01\)00157-3](https://doi.org/10.1016/s0010-0277(01)00157-3)


- McGurk, H. & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264, 746–748. <https://doi.org/10.1038/264746a0>
- McMurray, B., Kovack-Lesh, K. A., Goodwin, D. & McEchron, W. (2013). Infant-directed speech and the development of speech perception: Enhancing development or unintended consequence? *Cognition*, 129, 362–378. <https://doi.org/10.1016/j.cognition.2013.07.015>
- Mehler, J., Jusczyk, P. W., Lambertz, G., Halsted, N., Bertoncini, J. & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, 29, 143–178. [https://doi.org/10.1016/0010-0277\(88\)90035-2](https://doi.org/10.1016/0010-0277(88)90035-2)
- Metsala, J. L. & Walley, A. C. (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and early reading ability. In L. C. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 89–120). Hillsdale, NJ: Erlbaum
- Michaelis, R. (2010). Entwicklungsneurologie. In R. Michaelis & G. Niemann (Hrsg.), *Entwicklungsneurologie und Neuropädiatrie. Grundlagen und diagnostische Strategien* (S. 5–147). Stuttgart: Thieme.
- Moon, C., Cooper, R. P. & Fifer, W. P. (1993). Two-day-olds prefer their native language. *Infant Behavior & Development*, 16, 495–500. [https://doi.org/10.1016/0163-6383\(93\)80007-U](https://doi.org/10.1016/0163-6383(93)80007-U)
- Moon, C., Lagercrantz, H. & Kuhl, P. K. (2013). Language experienced in utero affects vowel perception after birth: A two-country study. *Acta Paediatrica*, 102, 156–160. <https://doi.org/10.1111/apa.12098>
- Morgan, L. & Wren, Y. E. (2018). A systematic review of the literature on early vocalizations and babbling patterns in young children. *Communication Disorders Quarterly*, 40, 3–14. <https://doi.org/10.1177/15257401187602015>
- Narayan, C. R., Werker, J. F. & Speeter Beddor, P. (2010). The interaction between acoustic salience and language experience in developmental speech perception: Evidence from nasal place discrimination. *Developmental Science*, 13, 407–420. <https://doi.org/10.1111/cdev.12087>
- Nespor, M., Pena, M. & Mehler, J. (2003). On the different role of vowels and consonants in speech processing and language acquisition. *Lingua e Linguaggio*, 2, 203–229. <https://doi.org/10.1418/10879>
- Papoušek, M. (2007). Communication in early infancy: An arena of intersubjective learning. *Infant Behavior and Development*, 30, 258–266. <https://doi.org/10.1016/j.infbeh.2007.02.003>
- Peña, M., Maki, A., Kovacic, D., Dehaene-Lambertz, G., Koizumi, H., Bouquet, F. & Mehler, J. (2003). Sounds and silence: An optical topography study of language recognition at birth. *Proceedings of the National Academy of Sciences*, 100, 11702–11705. <https://doi.org/10.1073/pnas.1934290100>
- Peterson, G. E. & Barney, H. L. (1952). Control methods used in a study of vowels. *Journal of the Acoustical Society of America*, 24, 175–184. <https://doi.org/10.1121/1.1906875>
- Petitto, L. A., Berens, M. S., Kovelman, I., Dubins, M. H., Jasinska, K. & Shalinsky, M. (2012). The “Perceptual Wedge Hypothesis” as the basis for bilingual babies phonetic processing advantage: New insights from fNIRS brain imaging. *Brain and Language*, 121(2), 142–155. <https://doi.org/10.1016/j.bandl.2011.05.003>
- Polka, L. & Werker, J. F. (1994). Developmental changes in the perception on non-native vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance*, 20, 421–435. <https://doi.org/10.1037/0096-1523.20.2.421>
- Pons, F., Lewkowicz, D. J., Soto-Faraco, S. & Sebastián-Gallés, N. (2009). Narrowing of intersensory speech perception in infancy. *Proceedings of the National Academy of Sciences*, 106, 10598–10602. <https://doi.org/10.1073/pnas.0904134106>
- Pons, F., Bosch, L. & Lewkowicz, D. J. (2015). Bilingualism modulates infants’ selective attention to the mouth of a talking face. *Psychological Science*, 26(4), 490–498. <https://doi.org/10.1177/0956797614568>
- Sansavini, A., Bertoncini, J. & Giovanelli, G. (1997). Newborns discriminate the rhythm of multisyllabic stressed words. *Development Psychology*, 66, 3–11. <https://doi.org/10.1037/0012-1649.33.1.3>
- Schaadt, G., Werwach, A., Obrig, H., Friederici, A. D. & Männel, C. (2023). Maturation of consonant perception, but not vowel perception, predicts lexical skills at 12 months. *Child Development*, 94(3), e166–e180. <https://doi.org/10.1111/cdev.13892>
- Sebastian-Galles, N. (2010). Bilingual language acquisition: Where does the difference lie? *Human Development*, 53, 245–255. <https://doi.org/10.1159/000321282>
- Sebastian-Galles, N., Albareda, B., Weikum, W. & Werker, J. F. (2012). A bilingual advantage in visual language discrimination in infancy. *Psychological Science*, 23(9), 994–999. <https://doi.org/10.1177/09567976124368>
- Shaw, K. E. & Bortfeld, H. (2015). Sources of confusion in infant audiovisual speech perception research. *Frontiers in Psychology*, 6, 1844. <https://doi.org/10.3389/fpsyg.2015.01844>
- Shi, R., Werker, J. F. & Morgan, J. L. (1999). Newborn infants’ sensitivity to perceptual cues to lexical and grammatical words. *Cognition*, 72, B11–B21. <https://doi.org/10.1016/s0010-0277>
- Singh, L., Loh, D. & Xia, N. G. (2017). Bilingual infants demonstrate perceptual flexibility in phoneme discrimination but perceptual constraint in face discrimination. *Frontiers in Psychology*, 8, 1563. <https://doi.org/10.3389/fpsyg.2017.01563>
- Skeide, M. A. & Friederici, A. F. (2016). The ontogeny of the cortical language network. *Nature Review Neuroscience*, 17, 323–332. <https://doi.org/10.1038/nrn.2016.23>
- Stager, C. L. & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature*, 388, 381–382. <https://doi.org/10.1038/41102>
- Streri, A., Coulon, M., Marie, J. & Yeung, H. H. (2016). Developmental change in infants’ detection of visual faces that match auditory vowels. *Infancy*, 21(2), 177–198. <https://doi.org/10.1111/inf.12104>
- Sundara, M., Polka, L. & Molnar, M. (2008). Development of coronal stop perception: Bilingual infants keep pace with their monolingual peers. *Cognition*, 108, 232–242. <https://doi.org/10.1016/j.cognition.2007.12.013>
- Swingle, D. (2009). Contributions of infant word learning to language development. *Philosophical Transactions of the Royal Society B*, 364, 3617–3632. <https://doi.org/10.1098/rstb.2009.0107>
- Szagan, G. (2019). *Sprachentwicklung beim Kind. Ein Lehrbuch*. Weinheim: Beltz.
- Tallal, P. (1980). Auditory temporal perception, phonics, and reading disabilities in children. *Brain and Language*, 9, 182–198. [https://doi.org/10.1016/0093-934X\(80\)90139-X](https://doi.org/10.1016/0093-934X(80)90139-X)
- Teinonen, T., Aslin, R. N., Alku, P. & Csibra, G. (2008). Visual speech contributes to phonetic learning in 6-month-old infants. *Cognition*, 108, 850–855. <https://doi.org/10.1016/j.cognition.2008.05.009>
- Thiessen, E. D. (2011). When variability matters more than meaning: The effect of lexical forms on use of phonemic contrasts. *Developmental Psychology*, 47, 1448–1458. <https://doi.org/10.1037/a0024439>
- Thiessen, E. D. & Pavlik Jr, P. I. (2016). Modeling the role of distributional information in children’s use of phonemic contrasts. *Journal of Memory and Language*, 88, 117–132. <https://doi.org/10.1016/j.jml.2016.01.003>
- Tincoff, R. & Jusczyk, P. W. (2012). Six-month-olds comprehend words that refer to parts of the body. *Infancy*, 17, 432–444. <https://doi.org/10.1111/j.1532-7078.2011.00084.x>
- Tomalski, P., Ribeiro, H., Ballieux, H., Axelsson, E. L., Murphy, E., Moore, D. G. & Kushnerenko, E. (2013). Exploring early develop-

- mental changes in face scanning patterns during the perception of audiovisual mismatch of speech cues. *European Journal of Developmental Psychology*, 10(5), 611–624. <https://doi.org/10.1080/17405629.2012.728076>
- Tsao F.-M., Liu H.-M. & Kuhl P. K. (2004). Speech perception in infancy predicts language development in the second year of life: A longitudinal study. *Child Development*, 75, 1067–1084. <https://doi.org/10.1111/j.1467-8624.2004.00726.x>
- Tsuji, S. & Cristia, A. (2014). Perceptual attunement in vowels: A meta-analysis. *Developmental Psychobiology*, 56, 179–191. <https://doi.org/10.1002/dev.21179>
- Vandewalle, E., Boets, B., Ghesquiere, P. & Zink, I. (2012). Auditory processing and speech perception in children with specific language impairment: Relations with oral language and literacy skills. *Research in Developmental Disabilities*, 33, 635–644. <https://doi.org/10.1016/j.ridd.2011.11.005>
- Vandermosten, M., Boets, B., Luts, H., Poelmans, H., Golestani, N., Wouters, J. & Ghesquiere, P. (2010). Adults with dyslexia are impaired in categorizing speech and nonspeech sounds on the basis of temporal cues. *Proceedings of the National Academy of Sciences*, 107, 10389–10394. <https://doi.org/10.1073/pnas.0912858107>
- Vihman, M. M. (1996). *Phonological development. The origins of language in the child*. Oxford, UK: Blackwell.
- Vouloumanos, A. & Werker, J. F. (2007). Listening to language at birth: Evidence for a bias for speech in neonates. *Developmental Science*, 10, 159–164. <https://doi.org/10.1111/j.1467-7687.2007.00549.x>
- Weikum, W., Vouloumanos, A., Navarra, J., Soto-Faraco, S., Sebastian-Galles, N. & Werker, J. F. (2007). Visual language discrimination in infancy. *Science*, 316(5828), 1159. <https://doi.org/10.1126/science.113768>
- Werker, J. F. & Tees, R. C. (1999). Influences on infant speech processing: Toward a new synthesis. *Annual Review of Psychology*, 50, 509–535. <https://doi.org/10.1146/annurev.psych.50.1.509>
- Werker, J. F. & Curtin, S. (2005). PRIMIR: A developmental framework of infant speech processing. *Language Learning and Development*, 1, 197–234. [https://doi.org/10.1207/s15473341lld0102\\_4](https://doi.org/10.1207/s15473341lld0102_4)
- Werker, J. F. & Hensch, T. K. (2015). Critical periods in speech perception. New directions. *Annual Review of Psychology*, 66, 173–196. <https://doi.org/10.1146/annurev-psych-010814-015104>
- Westermann, G. & Miranda, E. R. (2004). A new model of sensorimotor coupling in the development of speech. *Brain and Language*, 89, 393–400. [https://doi.org/10.1016/S0093-934X\(03\)00345-6](https://doi.org/10.1016/S0093-934X(03)00345-6)
- Yeung, H. H. & Werker, J. F. (2009). Learning words' sounds before learning how words sound: 9-month-olds use distinct objects as cues to categorize speech information. *Cognition*, 113, 234–243. <https://doi.org/10.1016/j.cognition.2009.08.010>
- Yeung, H. H. & Werker, J. F. (2013). Lip movements affect infants' audiovisual speech perception. *Psychological Science*, 24(5), 603–612. <https://doi.org/10.1177/0956797612458802>
- Zangl, R. & Mills, D. L. (2007). Increased brain activity to infant-directed speech in 6- and 13-month-old infants. *Infancy*, 11(1), 31–62. [https://doi.org/10.1207/s15327078in1101\\_2](https://doi.org/10.1207/s15327078in1101_2)
- Zimmer, E. Z., Fifer, W. P., Kim, Y. I., Rey, H. R., Chao, C. R. & Myers, M. M. (1993). Response of the premature fetus to stimulation by speech sounds. *Early Human Development*, 33, 207–215. [https://doi.org/10.1016/0378-3782\(93\)90147-M](https://doi.org/10.1016/0378-3782(93)90147-M)

#### Förderung

Open Access-Veröffentlichung ermöglicht durch die Universität Mannheim.

#### ORCID

Nicole Altvater-Mackensen  
 <https://orcid.org/0000-0002-8075-4720>

#### Prof. Dr. Nicole Altvater-Mackensen

Anglistik I/ Psycholinguistik  
 Universität Mannheim  
 Schloss  
 68161 Mannheim  
 Deutschland  
[altvater-mackensen@uni-mannheim.de](mailto:altvater-mackensen@uni-mannheim.de)