

*Analyse melodischer und artikulatorischer
Eigenschaften in Komfortvokalisationen
von Säuglingen mit deutscher
Muttersprache im Alter von 3-7 Monaten*

Inaugural-Dissertation

zur Erlangung der Doktorwürde
der Fakultät für Humanwissenschaften
der Julius-Maximilians-Universität Würzburg

Vorgelegt von
Verena Frank
aus Würzburg

Würzburg, 2021



Erstgutachter: Professor Dr. D. Hansen

Zweitgutachterin: Professor Dr. K. Wermke

Tag des Kolloquiums: 16.02.2022

Inhalt

Abbildungsverzeichnis	6
Tabellenverzeichnis	8
I. Einleitung	10
II. Theoretische Grundlagen	13
1. Anatomische und neurophysiologische Entwicklung der an der Lautproduktion beteiligten Strukturen und Mechanismen zwischen dem 3. und 7. Monat.....	13
1.1 Anatomische Reifung der laryngealen Ebene (Phonationssystem).....	13
1.2 Anatomische Reifung der supralaryngealen Ebene (Artikulationssystem)	15
1.3 Neurophysiologische Reifung und motorische Kontrolle der laryngealen und supralaryngealen Ebene	17
2. Die Entwicklung produktiver Fähigkeiten zwischen dem 3. und 7. Monat.....	23
2.1 Kategorisierung der Lautproduktionen auf Grundlage phasenbasierter Modelle	23
2.2 Akustische Eigenschaften von Komfortvokalisationen zwischen dem 3. und 7. Monat.....	30
2.3 Artikulatorische Fähigkeiten zwischen dem 3. und 7. Monat.....	33
3. Methodologie und Forschungsfragen	35
4. Material und Methoden	46
4.1 Probanden.....	46
4.2 Datenerhebung - Lautaufnahmen	47
4.3 Datenaufbereitung.....	49
4.4 Berechnung der Grundfrequenz- und Intensitätsverläufe.....	51
4.5 Spezifische Datenanalyse	51
4.5.1 Identifizierung von Komfortvokalisationen	52
4.5.2 Melodieanalyse	54
4.5.3 Melodiestrukturanalyse	56
4.5.4 Ermittlung des Melodiekomplexitätsgrades der Vokalisationen (MCI)	60
4.5.5 Quantitative Analyse – Erfassung akustischer Kenngrößen der Melodie..	61
4.5.6 Analyse artikulatorischer Fähigkeiten.....	63
4.5.6.1 Erfassung artikulatorischer Elemente pro Vokalisation	63
4.5.6.2 Analyse von Resonanzveränderungen	65

4.5.6.3 Analyse konsonanten-ähnlicher Elemente.....	66
4.6 Statistische Analyse	72
5. Ergebnisse	73
5.1 Charakterisierung akustischer Kenngrößen von Komfort-Vokalisationen im Alter von 3-7 Monaten	73
5.1.1 Mittlere Grundfrequenz und F_0 -Range (Hub).....	73
5.1.1.1 Lautbasierte Analyse der mittleren Grundfrequenz und des F_0 - Ranges	73
5.1.1.2 Kindbasierte Analyse der mittleren Grundfrequenz und des F_0 - Ranges	84
5.1.2 Analyse der Vokalisationslänge	88
5.1.2.1 Lautbasierte Analyse der Vokalisationslänge	88
5.1.2.2 Kindbasierte Analyse der Vokalisationslänge	93
5.2 Charakterisierung der Melodie-Struktur	95
5.3 Individuelle Variation der Melodiestructur - der Melodiekomplexitätsindex (MCI)	98
5.4 Charakterisierung der artikulatorischen Fähigkeiten	102
5.4.1 Die relative Häufigkeit artikulierter Vokalisationen.....	102
5.4.2 Die Komplexität der Artikulation - Anzahl artikulatorischer Elemente pro Vokalisation	103
5.4.3 Die Produktion artikulatorischer Elemente	106
5.4.3.1 Vokalisationsbasierte Analyse artikulatorischer Elemente	106
5.4.3.2 Kindbasierte Analyse artikulatorischer Elemente	109
5.4.4 Produktion artikulatorischer Elemente in Abhängigkeit von der Melodiestructur	112
6. Diskussion	114
6.1 Akustische Eigenschaften in Komfortvokalisationen im Alter von 3-7 Monaten	114
6.1.1 Grundfrequenzeigenschaften	115
6.1.2 Die Veränderung der Vokalisationslänge im Untersuchungszeitraum	122
6.2 Prosodierelevante Eigenschaften in Komfortvokalisationen im Alter von 3-7 Monaten	125

6.3 Artikulatorische Eigenschaften in Komfortvokalisationen im Alter von 3-7 Monaten	131
6.4 Zusammenspiel von Komplexität in Artikulation und Phonation in Komfortvokalisationen im Alter von 3-7 Monaten.....	135
III Conclusio	137
Literaturverzeichnis	144
Anhang	158
Danksagung.....	165

Abbildungsverzeichnis

Abbildung 1: Schmalbandspektrogramm (unteres Fenster) mit korrespondierender Zeitbereichsdarstellung (oberes Fenster) einer Einzelvokalisation	50
Abbildung 2: Zusammensetzung des Vokalisations-Repertoires	52
Abbildung 3: Beispiel für ein Melodie-Intensitätsdiagramm einer Vokalisation im CDAP-Programm. Die Dimensionsangaben erfolgen programmbedingt englisch.....	55
Abbildung 4: Vermessungsfenster im CDAP-Programm.....	62
Abbildung 5: Breitbandspektrogramm (unteres Fenster) mit Amplitudenverlauf (oberes Fenster) im Ausgabefenster der <i>PRAAT</i> -Software.	64
Abbildung 6: Kodierungsschema von Artikulationsbewegungen innerhalb einer Vokalisation.....	65
Abbildung 7: Darstellung einer Resonanzveränderung (rv) innerhalb einer Vokalisation im Schmalbandspektrogramm im Ausgabefenster der <i>PRAAT</i> -Software.	66
Abbildung 8: Artikulationsbereiche und -organe des Mundraums (Weinrich & Zehner, 2017), S.7.	67
Abbildung 9: Boxplot (oben) und Histogramm (unten) der mittleren F_0 -Daten im Gesamtzeitraum.....	75
Abbildung 10: Boxplot (oben) und Histogramm (unten) des absoluten F_0 -Hubes im Gesamtzeitraum.....	77
Abbildung 11: Boxplot der mittleren F_0 für die einzelnen Monate im Untersuchungszeitraum mit eingezeichnetem Gruppenmedian (graue Linie)	79
Abbildung 12: Boxplot für den absoluten F_0 -Hub innerhalb einer Vokalisation pro Monat mit eingezeichnetem Gruppenmedian (graue Linie).....	80
Abbildung 13: Boxplot der mittleren F_0 für die einzelnen Kinder mit eingezeichnetem Gruppenmedian	82
Abbildung 14: Boxplot für den absoluten F_0 -Hub [Hz] pro Kind mit eingezeichnetem Gruppenmedian (graue Linie).....	83
Abbildung 15: Boxplot für die Veränderung der mittleren F_0 [Hz] im Untersuchungszeitraum (kindbasiert)	85
Abbildung 16: Boxplot für den absoluten F_0 -Hub pro Monat (kindbasiert)	86
Abbildung 17: Boxplot für den relativen F_0 -Hub (kindbasiert) mit horizontalen Orientierungslinien bei der musikalischen Quinte (1,49) und Oktave (2,0).....	87
Abbildung 18: Boxplot (oben) und Histogramm (unten) der Vokalisationslänge [ms] im Gesamtzeitraum.....	89

Abbildung 19: Boxplot für die Vokalisationslänge pro Monat	91
Abbildung 20: Liniendiagramm für den Median der Lautlänge [ms] pro Monat mit eingezeichneten 95%-Konfidenzintervallen (graue Whisker).....	91
Abbildung 21: Boxplot der Vokalisationslänge [ms] pro Kind mit eingezeichnetem Gruppenmedian (graue Linie).....	93
Abbildung 22: Boxplot für die mittlere Vokalisationslänge pro Monat (kindbasiert).....	94
Abbildung 23: Verteilung der zusammengefassten Melodiestrukturen (Einfachbögen+KL, Mehrfachbögen, segmentierte Laute) pro Monat.....	97
Abbildung 24: Verteilung der MCI-Werte pro Kind (es gehen alle Aufnahmetage pro Kind ein) mit eingezeichnetem Gruppenmedian (graue Linie)	100
Abbildung 25: Boxplot für den MCI pro Monat (pro Monat geht von jedem Kind ein Wert ein)	101
Abbildung 26: Kreisdiagramm für die Häufigkeitsverteilung der Laute mit einem, zwei sowie drei und mehr artikulatorischen Elementen im Gesamtzeitraum	104
Abbildung 27: Balkendiagramm für die Häufigkeitsverteilung der Vokalisierungen mit einem, zwei sowie drei und mehr artikulatorischen Elementen pro Monat	106
Abbildung 28: Kreisdiagramm für die Häufigkeitsverteilung einzelner artikulatorischer Elemente im Untersuchungszeitraum.....	107
Abbildung 29: Balkendiagramm für die Häufigkeiten einzelner artikulatorischer Elemente pro Monat.....	109
Abbildung 30: Boxplot für die Häufigkeitsverteilung einzelner artikulatorischer Elemente (kindgemittelt) im Untersuchungszeitraum (Gesamtprozentwerte der einzelnen Kinder gehen ein).....	111
Abbildung 31: Liniendiagramm für die Verteilung artikulierter Vokalisierungen über die verschiedenen Melodiestrukturen pro Monat (Strukturen mit einem Bogen, Strukturen mit mehreren Bögen, segmentierte Strukturen)	112

Tabellenverzeichnis

Tabelle 1: Probandentabelle mit Aufnahmezeitpunkten pro Kind	49
Tabelle 2: Anzahl der analysierten Vokalisationen pro Kind/pro Monat.....	53
Tabelle 3: Strukturkategorien zur Analyse der Melodiestructur (Wermke, 2002; Wermke et al., 2007)	57
Tabelle 4: Kodierung der Artikulationsbewegungen	68
Tabelle 5: Deskriptive Statistik der mittleren F_0 im Gesamtzeitraum	74
Tabelle 6: Deskriptive Statistik von absolutem und relativem F_0 -Hub im Gesamtzeitraum	76
Tabelle 7: Deskriptive Statistik der mittleren F_0 pro Monat	78
Tabelle 8: Deskriptive Statistik des absoluten F_0 -Hubes pro Monat	80
Tabelle 9: Deskriptive Statistik der mittleren F_0 pro Kind im Untersuchungszeitraum ..	81
Tabelle 10: Deskriptive Statistik für den absoluten F_0 -Hub pro Kind	83
Tabelle 11: Deskriptive Statistik der mittleren F_0 pro Monat (kindbasierte Mittelung) ..	84
Tabelle 12: Deskriptive Statistik des absoluten F_0 -Hubes pro Monat (kindbasiert)	85
Tabelle 13: Deskriptive Statistik relativer F_0 -Hub (kindbasierte Mittelung).....	87
Tabelle 14: Deskriptive Statistik der Lautlänge	88
Tabelle 15: Deskriptive Statistik der Lautlänge [ms] pro Monat.....	90
Tabelle 16: Deskriptive Statistik der mittleren Vokalisationslänge pro Kind.....	92
Tabelle 17: Deskriptive Statistik für die mittlere Vokalisationslänge pro Monat (kindbasierte Mittelung)	94
Tabelle 18: Häufigkeiten einzelner Melodie-Struktur-Typen im Untersuchungszeitraum	95
Tabelle 19: Auftrittshäufigkeiten für einfache, komplexe und segmentierte Lautstrukturen pro Monat	96
Tabelle 20: Deskriptive Statistik für den MCI (alle Aufnahmetage gehen ein)	98
Tabelle 21: Deskriptive Statistik für den MCI pro Kind (alle Aufnahmetage gehen ein)	99
Tabelle 22: MCI pro Monat (ein Monatswert pro Kind)	101
Tabelle 23: Häufigkeiten der Vokalisationen mit und ohne Artikulation pro Monat und Gesamtzeitraum	103
Tabelle 24: Häufigkeitsverteilung der Vokalisationen mit einem, zwei sowie drei und mehr artikulatorischen Elementen pro Monat und Gesamtzeitraum.....	105

Tabelle 25: Häufigkeitsverteilung unterschiedlicher artikulatorischer Elemente im Untersuchungszeitraum	107
Tabelle 26: Anzahl und Häufigkeiten einzelner artikulatorischer Elemente pro Monat	108
Tabelle 27: Häufigkeitsverteilung einzelner artikulatorischer Elemente (kindgemittelt) im Untersuchungszeitraum (pro Kind geht ein Gesamtprozentwert pro Artikulationsort ein)	110
Tabelle 28: Relative Häufigkeit einzelner artikulatorischer Elemente pro Kind im Untersuchungszeitraum [%]	111
Tabelle 29: Artikulierte Vokalisationen in Abhängigkeit von der Melodiestructur pro Monat (grau hinterlegte Felder)	113
Tabelle 30: Eckdaten der Studien zu akustischen und phonetischen Eigenschaften in Komfortvokalisationen im Alter zwischen 3 und 7 Monaten	158
Tabelle 31: Ergebnisse der standardisierten Testverfahren zur sprachlichen Entwicklung. (ua: unauffällig; a: auffällig, rot unterlegt)	164

I. Einleitung

„[...] Man has an instinctive tendency to speak, as we see in the babble of our young children [...]“.

Charles Darwin, Decent of Man, 1871, p. 53

Der Mensch hat eine angeborene Fähigkeit, Sprache zu erlernen. So ist es unbestritten, dass Kinder mit einem angeborenem Entwicklungsprogramm ausgestattet sind, welches sie befähigt, Sprache zu erwerben. Anatomische und sozial-kognitive Reifungsvorgänge sowie Lernprozesse in sozialer Interaktion sind Voraussetzung zur Entfaltung dieses angeborenem Entwicklungsprogramms (Ibbotson & Tomasello, 2017; Yang, Crain, Berwick, Chomsky, & Bolhuis, 2017). Säuglinge speichern und erwerben schon in den ersten Lebensmonaten bzw. teilweise sogar vor der Geburt prosodische Informationen über ihre Muttersprache. Wesentliche Meilensteine der Sprachentwicklung in den ersten fünf Lebensjahren werden von allen Kindern in der gleichen Reihenfolge durchlaufen. Die wichtigsten Grundbausteine für den Spracherwerb werden in den ersten drei Lebensjahren gesetzt. Die spracherwerbsrelevante Entwicklung beginnt bereits auf dem Weg zum ersten Wort, also während der Phase der „vorsprachlichen“ Entwicklung im ersten Lebensjahr (Grimm & Weinert, 2002; Sachse, 2007).

Sprachliche Fähigkeiten gelten als maßgeblich für den späteren Bildungserfolg und stellen zweifelsohne eine entscheidende Schlüsselkompetenz in unserem formalen Bildungssystem dar. Demzufolge stehen seit einigen Jahren v.a. auch frühe sprachliche Erwerbsprozesse im Fokus der Spracherwerbsforschung (Sachse et al. 2020).

Der Ausbau des Betreuungsangebots von Kindern unter drei Jahren hat u.a. dazu geführt, dass auch die Frühpädagogik und Politik mittlerweile auf diesen Zusammenhang aufmerksam wurde und erste Programme und Initiativen startete, um Spracherwerbsprozesse im frühpädagogischen Setting zu untersuchen und durch Fachkräfte professionell zu unterstützen und zu begleiten (z.B. Bundesprogramm „Frühe Chancen“¹, Bundesprogramm „Sprach-Kitas: Weil Sprache der Schlüssel zur Welt ist“²). Neben der Unterstützung früher Spracherwerbsprozesse durch Angebote der alltagsintegrierten sprachlichen Bildung im Rahmen von Betreuungsangeboten, z.B. in Kindertagesstätten, sind aus sprachheilpädagogischer Sicht v.a. Erwerbsverläufe von Risikogruppen, also Kindern mit besonderem Förderbedarf, relevant. Hierzu zählen u.a.

¹ Nähere Informationen unter <https://www.fruehe-chancen.de/>

² Nähere Informationen unter <https://sprach-kitas.fruehe-chancen.de/>

Kinder mit Hörstörungen, mehrsprachig aufwachsende Kinder oder sog. „Late Talker“, also Kinder, die mit 24 Monaten noch nicht in die produktive Sprache eingestiegen sind. Diese Entwicklung ist aus sprachheilpädagogischer Sicht sehr positiv zu sehen, ist allerdings trotz umfangreicher Forschungen aufgrund von fehlenden Längsschnittuntersuchungen noch ausbaufähig (Bartl-Pokorny et al., 2013; Egert, Sachse, & Groth, 2016; Grossheinrich et al., 2019).

Um geeignete Methoden zur Früherkennung und Frühförderung zu entwickeln, ist es zunächst erforderlich, die ungestörten vorsprachlichen Entwicklungsverläufe von sich normal entwickelnden Kindern systematisch zu untersuchen. Dazu sind u.a. Längsschnittuntersuchungen zu akustischen und phonetischen Eigenschaften der Vokalisationen dieser frühen Phasen geeignet. Hier setzt die vorliegende Arbeit an. Die Phase der Schreivokalisationen sowie die Phase zielsprachlich geprägter Silbenbabblers (Kanonisches Babbeln) wurden am ZVES bereits ausführlich untersucht (Pachtner, 2016; Wermke, 2007). Während die Arbeit von Pachtner (2016) erstmalig Referenzwerte zu akustischen und phonetischen Eigenschaften in der Phase des Kanonischen Babbelns vorlegte, soll diese Arbeit nun erstmalig Referenzwerte für akustische und artikulatorische Eigenschaften von Komfortvokalisationen bei deutsch erwerbenden Kindern im Alter zwischen 3 und 7 Monaten vorlegen und damit die Lücke zwischen Schreivokalisationen und Kanonischem Babbeln schließen.

Die Besonderheit dieser Phase liegt, bedingt durch anatomische, neurophysiologische und sozial-kognitive Reifungsvorgänge, im Zusammenschalten von zwei produktiven Mechanismen der menschlichen Sprache: Phonation und Artikulation. Ziel der vorliegenden explorativen Arbeit war es, diese beiden am Anfang der Entwicklung noch „getrennten“ Systeme objektiv zu untersuchen.

Bisher gibt es bereits einige, in der Spracherwerbsforschung verbreitete Modelle zum Ablauf der vorsprachlichen Entwicklung für andere Umgebungssprachen (Koopmans-van Beinum & van der Stelt, 1986; Oller, 1980; Stark, 1980). Diese Modelle beschreiben in erster Linie den zeitlichen Ablauf vorsprachlicher Phasen und die darin charakteristisch auftretenden Vokalisationstypen. Auch wenn diese Modelle, die häufig auf Beobachtungen von nur sehr wenigen Probanden basieren, einen wichtigen Beitrag zur Interpretation möglicher Entwicklungszusammenhänge und dem Ablauf der vorsprachlichen Entwicklung liefern, fehlen bisher jedoch objektive und systematische Längsschnittuntersuchungen zur Beschreibung quantitativer und qualitativer Lauteigenschaften dieser frühen Erwerbsphasen.

Als theoretische Basis vorliegender Arbeit gilt, wie in den genannten Modellen auch, die Annahme einer kontinuierlich ablaufenden vorsprachlichen Entwicklung hin zur frühen

Sprachentwicklung im engeren (linguistischen) Sinne, die sog. Kontinuitätshypothese. Sowohl die Entwicklung der frühen Schreivokalisationen als auch die der Komfortvokalisationen durchläuft eine Folge von Entwicklungsstufen von zunächst einfachen hin zu zunehmend komplexeren Strukturen (Stoel-Gammon, 1998; Wermke & Mende, 2011). „All infants, regardless of the linguistic community in which they are being raised, pass through the same stages of vocal development “(Stoel-Gammon, 1998, p. 89).

Hierbei wird eine Entwicklungsreihenfolge von der suprasegmentalen zur segmentalen Ebene angenommen. Während in den ersten Lebensmonaten die Entwicklung prosodischer Elemente (Melodie, Rhythmus) in Form von laryngealen Mechanismen im Vordergrund steht, erlangt ab dem 3. Lebensmonat die segmentale Ebene in Form von ersten supralaryngeal produzierten artikulatorischen Elementen zunehmende Bedeutung (ebd.). Diese Entwicklungsabfolge wurde auch in der vorliegenden Arbeit betrachtet. So wurde die Melodie als Grundbaustein und übergeordnete Struktur von Komfortvokalisationen in Zusammenhang mit frühen supralaryngealen Fähigkeiten untersucht.

Kapitel II stellt den theoretischen Teil der vorliegenden Arbeit dar. Er beinhaltet die Kapitel 1-5. In Kapitel 1 wurde zunächst der aktuelle Forschungsstand zu anatomischen und neurophysiologischen Reifungsprozessen der an der Lautproduktion von Komfortvokalisationen beteiligten Strukturen und Mechanismen im Alter von 3-7 Monaten dargelegt. Kapitel 2 gibt einen Überblick über die Entwicklung produktiver Fähigkeiten. Im Einzelnen werden phasenbasierte Modelle zur vorsprachlichen Entwicklung vorgestellt und Studien zur Entwicklung akustischer und artikulatorischer Fähigkeiten in Komfortvokalisationen zusammengefasst. In den darauffolgenden beiden Kapiteln werden anhand methodologischer Überlegungen einzelne Forschungsfragen formuliert (Kapitel 3) und die konkrete methodische Herangehensweise erläutert (Kapitel 4). Kapitel 5 stellt jeweils die Ergebnisse der vokalisationsbasierten und kindbasierten Analysen dar und liefert damit erstmalig objektive Referenzwerte für Komfortvokalisationen deutsch erwerbender Kinder im Alter von 3-7 Monaten. Die Ergebnisse werden abschließend in Kapitel 6 unter den in Kapitel 1 und 2 dargestellten theoretischen Erkenntnissen diskutiert und mit den Forschungsfragen in Zusammenhang gebracht. Kapitel III beinhaltet eine Zusammenfassung und Vorschläge für die weitere Forschung.

II. Theoretische Grundlagen

1. Anatomische und neurophysiologische Entwicklung der an der Lautproduktion beteiligten Strukturen und Mechanismen zwischen dem 3. und 7. Monat

Im Laufe des ersten Lebensjahres vollziehen sich enorme anatomische und neurophysiologische Veränderungen der am Sprechakt beteiligten Systeme. Dazu gehören das Respirationssystem, der Larynx und der Vokaltrakt sowie die mit ihrer Funktion verbundenen neurophysiologischen Regelsysteme (Kent, 1984). Im Folgenden Kapitel werden zunächst Veränderungen der laryngealen Ebene zwischen dem dritten und siebten Monat beschrieben. Anschließend werden supralaryngeale anatomische Umstrukturierungen und neurophysiologische Reifungen des Vokaltrakts dargestellt. Die Veränderung dieser Strukturen hat einen großen Einfluss auf den Beginn und den Ablauf vorsprachlicher Entwicklungsstufen.

1.1 Anatomische Reifung der laryngealen Ebene (Phonationssystem)

Zum Zeitpunkt der Geburt ist der Larynx bereits wesentlicher Bestandteil wichtiger physiologischer Funktionen. Zu den wichtigsten zählen der Schutz der unteren Atemwege, die Regulation der respiratorischen Luftstromregelung, das Schlucken und die Phonation (Harding, 1984; Meyer, 2009; Tucker, 1993). Dennoch unterscheidet sich der Kehlkopf des Neugeborenen von dem des Erwachsenen und unterliegt in den ersten Lebensjahren funktionellen und morphologischen Veränderungen (Eckel et al., 2000; Kent & Vorperian, 1995; Lieberman D.E., McCarthy R.C., Hiiemae K.M., & Palmer J.B., 2001).

Beim Neugeborenen liegt der Kehlkopf in Relation zur Wirbelsäule wesentlich höher als beim Erwachsenen. Während er beim Säugling noch zwischen dem zweiten und vierten Halswirbel lokalisiert ist, senkt er sich bis in das Erwachsenenalter auf die Höhe zwischen dem fünften und siebten Halswirbel ab. Die relativ hohe Lage des infantilen Larynx

sichert den Schutz der unteren Atemwege während des Fütterns in Rückenlage. Die Absenkung des Larynx beginnt etwa zwischen dem vierten und sechsten Monat, wodurch sich die kindliche Form des Vokaltrakts langsam der rechtwinkligen Form des adulten Vokaltrakts annähert. In diesem Zeitraum entwickelt sich auch die Fähigkeit zur aufrechten Nahrungsaufnahme (Eckel et al., 2000; Kent, 1995; Tucker, 1993).

Der kindliche Kehlkopf unterscheidet sich nicht nur in Größe und Lage von dem des erwachsenen Organs. Die Größe der Glottis entspricht zum Zeitpunkt der Geburt etwa einem Drittel der Größe des Erwachsenen und unterliegt in den ersten 18 Monaten einem Wachstumsspur. Neben dem Größenunterschied unterscheidet sich auch die Oberflächenbeschaffenheit der Stimmlippen. Es vollzieht sich ein Wandel der Verhältnisse zwischen kartilaginären und membranösen Anteilen der Stimmlippen. So bestehen beim Neugeborenen ca. 50% der Stimmlippen aus einer kartilaginären Oberfläche. Dies ist auf den im Vergleich zum Erwachsenen relativ großen Aryknorpel zurückzuführen (Fried, Kelly J. H., & Strome, 1982, S. 561). Beim Erwachsenen hingegen besteht nur noch ein Drittel der Stimmlippenoberfläche aus Knorpel, während sich die anderen beiden Drittel aus Muskeln und anderem Weichteilgewebe (Muskosa, Submuskosa) zusammensetzen. Diese membranöse Stimmlippenoberfläche ist wesentlich biegsamer und flexibler als die des Neugeborenen (Eckel et al., 2000; Fried et al., 1982; Kent, 1995; Tucker, 1993).

Ein weiterer Unterschied auf glottaler Ebene besteht im histologischen Strukturaufbau der Stimmlippen. Die adulten Stimmlippen weisen eine mehrschichtige Struktur auf, welche sich aus dem Epithel (Deckgewebe), der Lamina propria (Oberflächenschicht, Zwischenschicht, Tiefenschicht) und dem Musc. vocalis zusammensetzt. Letzterer formt den Hauptkörper der Stimmlippen (Gray, Hirano, & Sato, 1993; Sato, Hirano, & Nakashima, 2001). Die obere Schicht der Lamina propria (Reinke Raum) besteht aus lockerem und leicht verschiebbarem Gewebe und stellt die Schicht dar, welche während der Phonation am stärksten vibriert. Die mittlere Schicht, welche wesentlich „steifer“ als die obere Schicht ist, setzt sich aus Elastinfasern zusammen, während die tiefste Schicht aus Kollagenfasern besteht. Die Zwischenschicht und Tiefenschicht bilden zusammen das Stimmband (lat. Ligamentum vocale) (Gray et al., 1993). Mechanisch lassen sich diese fünf Schichten durch die *cover-body theory of phonation* (Hirano, 1975) beschreiben. „Cover“ umfasst das Epithel und die elastene Oberflächenschicht der Lamina propria. Als „body“ definiert Hirano (1975) den Musc. vocalis. Die Zwischenebene, welche das Stimmband darstellt, beschreibt er als „transition“ (zit. n. Boseley & Hartnick, 2006; Gray et al., 1993; Hartnick, Rehbar, & Prasad, 2005). Beim Neugeborenen besitzt die Lamina propria noch keine mehrschichtige Struktur. Es gibt

keine Anzeichen für die Ausprägung eines Stimmbands („transition“). Somit besteht die Lamina propria beim Neugeborenen allein aus einer relativ dicken hyperzellulären Monoschicht, welche strukturell dem Reinke Raum beim Erwachsenen entspricht (Gray et al., 1993; Hartnick et al., 2005; Sato et al., 2001). Trotz dieser unreifen Struktur ist sie bestens an die Bedürfnisse des Säuglings angepasst. So ist der Säugling im Vergleich zum Erwachsenen in der Lage über einen wesentlich längeren Zeitraum höhere Tonlagen und Intensitäten zu erzeugen (Säuglingsschrei) (Schweinfurth & Thibeault, 2008). Mit etwa zwei Monaten beginnt die Ausdifferenzierung in eine bilaminare Struktur mit unterschiedlichen, voneinander getrennten zellulären Populationsdichten (Hartnick et al., 2005). Diese Strukturveränderung könnte u.a. zeitlich mit dem vermehrten Auftreten von Quasivokalen und Gurrlauten - gekennzeichnet durch *normale Phonation*³ - in Zusammenhang stehen (Oller, 2000).

1.2 Anatomische Reifung der supralaryngealen Ebene (Artikulationssystem)

Neben den anatomisch-morphologischen Veränderungen auf laryngealer Ebene kommt es während der ersten sechs Lebensmonate auch zu einer anatomischen Restrukturierung im Bereich der supralaryngealen Ebene. Bei der Geburt ähnelt der Vokaltrakt des Säuglings eher dem eines nichtmenschlichen Primaten und entspricht daher nicht einfach einer Miniaturkopie des hominiden adulten Vokaltrakts (Bosma, 1975; Fletcher, 1973; Kent & Murray, 1982; Vorperian et al., 2005).

Ein wesentlicher Unterschied zwischen der infantilen und adulten Form des Vokaltraktes besteht in seiner Länge. Mit ca. 4 cm Länge, ist der Pharynx des Neugeborenen nur etwa ein Drittel so lang wie beim Erwachsenen, während die infantile Mundhöhle im Vergleich zur adulten Mundhöhle relativ breit ist. Weiterhin verläuft der oropharyngeale Kanal nicht wie beim Erwachsenen rechtwinkelig.

Die Zunge ist beim Neugeborenen im Verhältnis zum gesamten Vokaltrakt relativ groß und liegt in vorverlagerter Position. Durch die schwache Ausprägung der intrinsischen Zungenmuskulatur ist diese zunächst an starre Vor- und Rückwärtsbewegungen gebunden (Saugbewegung).

³ Der Begriff „normale“ oder auch „modale“ Phonation wird in der Literatur verwendet, um die klare ungestörte Phonation im für das Sprechen relevanten Grundfrequenzbereich zu charakterisieren (Titze, 1994; Oller, 2000).

Velopharynx und Epiglottis liegen beim Säugling sehr eng aneinander, wodurch der Säugling zum obligatorischen Nasenatmer wird, und in der Lage ist, während des Schluckens zu atmen (z.B. Kent, 1995; Kent & Murray, 1982).

In den ersten Lebensjahren unterliegt der kindliche Vokaltrakt einem drastischen Restrukturierungsprozess. Dieser beinhaltet anatomische Veränderungen in Form und Größe. Ausschlaggebend ist v.a. die Absenkung des Larynx (siehe 2.1) sowie des Zungenbeins und der Zunge. Mit ca. sechs Monaten ist die intrinsische Zungenmuskulatur kräftig genug auch vertikale Bewegungen auszuführen.

Hinzu kommt ein Längenwachstum des Vokaltraktes mit einer gleichzeitigen Verkleinerung des oro-laryngo-pharyngealen Winkels (Kent, 1995; Vorperian, Kent, Gentry, & Yandell, 1999). Aus Magnetresonanzstudien geht hervor, dass es bei den meisten dieser anatomischen Strukturen zu einem Wachstumsspurt zwischen dem ersten und vierten, sowie dem 13. und 15. Monat kommt (Vorperian et al., 1999; Vorperian et al., 2005).

Eine weitere entscheidende anatomische Veränderung ist das Lösen des velopharyngealen Kontakts durch die Entkopplung von Epiglottis und Velum sowie die Fähigkeit den Velopharynx durch das Velum zu schließen. Sie ist die Voraussetzung für eine Phonation ohne Hypernasalität und die Bildung von Verschlusslauten. Wann genau diese Fähigkeit erworben wird ist noch unklar (Thom, Hoit, Hixon, & Smith, 2006). Allerdings liefern Studien bezüglich des Zusammenhangs zwischen anatomischen Daten und akustischen Eigenschaften vorsprachlicher Vokalisationen indirekt Hinweise, die die Annahme eines velopharyngealen Verschlusses zwischen dem vierten und sechsten Monat indizieren (Kent & Murray, 1982; Oller, 1986; Robb, Chen, & Gilbert, 1997; Vorperian & Kent, 2007; Wermke, Mende, Manfredi, & Brusciagioni, 2002). Obwohl bereits ab dem zweiten Monat Verschlusslaute im hinteren Bereich des Vokaltraktes produziert werden können, geht man davon aus, dass der dafür benötigte velopharyngeale Mechanismus noch nicht stabil kontrolliert werden kann. Der weiche Gaumen (Velum) senkt sich erst mit zunehmender Reife in Relation zu seinem Ursprung in der Schädelbasis ab. Bis dahin funktioniert der Musc. levator palatini als Gaumensegelspanner. Nach der Absenkung des weichen Gaumens um den sechsten Lebensmonat übernimmt der Musc. levator palatini dann die Hebung des Velums. Diese anatomischen Veränderungen bilden die Grundlage für die Erzeugung nasal-oraler Kontraste und somit eine Erweiterung des Vokalisationsrepertoires (Fletcher, 1973; Thom et al., 2006).

Ogleich sich eine Vielzahl der respiratorischen, laryngealen und supralaryngealen Strukturen bis in das Erwachsenenalter weiterentwickeln, kann bereits an dieser Stelle festgehalten werden, dass sich die Strukturen des infantilen Vokaltrakts schon ab dem vierten Lebensmonat an die Strukturen des adulten Vokaltrakts angleichen. Welche spezifischen Auswirkungen diese anatomischen Veränderungen für die Phonation und Artikulation genau haben, ist noch nicht ausreichend bekannt. Es ist jedoch sicher, dass zumindest einigen von ihnen eine enorme Relevanz bei der Beurteilung vorsprachlicher Leistungen zukommt (Kent & Miolo, 1996).

1.3 Neurophysiologische Reifung und motorische Kontrolle der laryngealen und supralaryngealen Ebene

Die erhebliche Erweiterung des Vokalisationsrepertoires zwischen dem zweiten und siebten Monat ist neben den in Kapitel 1.2 beschriebenen anatomischen Veränderungen gleichzeitig auf eine zunehmende zentral-nervöse motorische Kontrolle bezüglich der an der Lautproduktion beteiligten Organe zurückzuführen (Kent, 1976). Die Produktion von Sprachlauten ist eine hoch komplexe und schnell ablaufende motorische Verhaltensweise, welche die Koordination von über 100 laryngealen, orofazialen und respiratorischen Muskeln erfordert. Auch die vorsprachlichen Lautäußerungen basieren bereits auf sehr schnellen neurophysiologischen Mechanismen (Simonyan & Horwitz, 2011). Auch wenn bereits viele Forschungsarbeiten zu den am Sprechvorgang beteiligten, neurophysiologischen Mechanismen Erwachsener veröffentlicht wurden, gibt es bezüglich detaillierter Kenntnisse über die Reifung der an der Lautproduktion beteiligten Strukturen im ersten Lebensjahr noch enormen Forschungsbedarf.

Im Folgenden sollen Reifungsverläufe grob anhand der Ergebnisse bisher durchgeführter Studien zur Organisation peripherer und zentraler Kontrollmechanismen abgeleitet werden. Diese basieren auf Untersuchungen (Neuro-Imaging- und Elektrostimulationsstudien) am Menschen und an nichtmenschlichen Primaten, welche ausführlich in einer Übersichtsarbeit von Simonyan und Horwitz (2011) zusammengefasst und diskutiert wurden. In dieser Übersichtsarbeit stellen sie die Hypothese auf, dass die Lage des laryngeal-motorischen Cortex (LMC) im primären motorischen Cortex und seine direkten Verbindungen zu den laryngealen Motoneuronen im Hirnstamm eine essenzielle Bedeutung in der menschlichen evolutionären Entwicklung hinsichtlich der Ausprägung sprachlicher Fähigkeiten, nämlich dem freiwilligen und willkürlichen Sprechen und Vokalisieren hatte (Simonyan & Horwitz, 2011).

Die zentrale Kontrolle der Stimm- und Sprachproduktion wird durch zwei parallele Leitungsbahnen im Gehirn ausgeführt. Zum einen durch die limbisch-vokale Leitungsbahn, die für die Kontrolle angeborener nonverbaler und emotionaler Vokalisationen zuständig ist. Zum anderen durch die laryngeal-kortikale Leitungsbahn, welche für die feinmotorische Kontrolle willkürlicher Stimmproduktion (wie z.B. beim Sprechen und Singen) sowie für die Produktion willkürlicher angeborener Vokalisationen verantwortlich ist. Beide Leitungsbahnen sind mit einer Basisebene im Rückenmark und im unteren Hirnstamm sowie einer komplexen Ebene im anterioren cingulären Cortex (ACC) und im laryngeal-motorischen Cortex (LMC) hierarchisch organisiert (Simonyan & Horwitz, 2011).

Hirnphysiologische Untersuchungen an Primaten lassen vermuten, dass sich die Kontrolle der Stimmproduktion beim menschlichen Säugling hierarchisch und in enger Abstimmung zu supralaryngealen Mechanismen entwickelt (z.B. Ploog, 1979).

Die unterste Ebene stellt die Ebene der sensorischen und motorischen Nervenkerne des unteren Hirnstamms (z.B. Nuclei ambiguus, trigeminus, facialis, hypoglossus, fasciculus solitarius) und des Rückenmarks dar. Sie ist für die basale Koordination der laryngealen, respiratorischen und artikulatorischen Muskelaktivität zuständig. Zusammen mit der *Formatio reticularis* (Brücke) und dem unteren Hirnstamm bilden sie beispielsweise wichtige Strukturen für die Stimmkontrolle während der Erzeugung angeborener Vokalisationen, wie z.B. dem Schmerzscrei und Lachen beim Säugling (Luthe, Hausler, & Jürgens, 2000). Die dorsal gelegenen Nervenkerne haben eine direkte Verbindung zu den laryngealen Motoneuronen im Nucleus ambiguus, den artikulatorischen Motoneuronen in den Nuclei trigeminus, facialis und hypoglossus, sowie zu den expiratorischen Motoneuronen der oberen Brust- und Lendenwirbelsäule.

Die neuronale Basis des Säuglingsschreis wurde lange Zeit mithilfe des sog. „Brainstem models“ beschrieben. Dabei ging man davon aus, dass das Vorder- und Mittelhirn in den ersten Lebensmonaten nicht an der Produktion angeborener Vokalisationen, wie bei nichtmenschlichen Primaten, beteiligt ist (Simonyan & Horwitz, 2011). Diese Ansicht ist allerdings auf Ergebnisse in Studien an jeweils einem anencephalen Neugeborenen zurückzuführen, welches in der Lage war, vokalisches auf Schmerzstimuli zu reagieren und nach heutiger Ansicht nicht mehr haltbar (z.B. Monnier & Willi, 1953; Newman, 2007)⁴.

⁴ Synonym wird diese Struktur in der Literatur auch mit dem Begriff „limbischer Cortex“ umschrieben, z.B. bei Newman (2007).

So spricht beispielsweise auch die Tatsache, dass angeborene Vokalisationen bereits in den ersten Lebensmonaten unter willkürlicher Kontrolle stattfinden können gegen diese veraltete Ansicht⁵. Die Tatsache, dass angeborene Vokalisationen willkürlich initiiert, und somit auch unter emotionaler Kontrolle produziert werden können, deutet auf eine höhere Ebene der Stimmkontrolle zu einem bereits relativ frühen Zeitpunkt in der Entwicklung hin. Dies ist durch einen Input aus Gehirnregionen wie dem periaquäduktalen Grau (PAG) und dem anterioren cingulären Cortex (ACC) gegeben. In neueren Magnetresonanztudien konnte gezeigt werden, dass die Myelinisierung dieser Strukturen bereits zwischen dem 3. und 4. Monat abgeschlossen ist (Deoni et al., 2011). Dies scheint die Annahme der funktionellen Beteiligung dieser Strukturen an frühesten willkürlich initiierten Vokalisationen zu bestätigen.

Das PAG spielt eine große Rolle beim Auslösen vokaler Reaktionen sowie der Modulation der Intensität dieser Reaktionen. Der ACC hingegen ist für den willkürlich kontrollierten Stimmeinsatz und seine emotionale Intonation zuständig. Die Formatio reticularis repräsentiert die grundlegende Ausführungsebene dieser Leitungsbahn (Simonyan & Horwitz, 2011). Auch diese Strukturen sind bereits bis zum vierten Monat myelinisiert, was die zunehmende Fähigkeit einer schnellen Frequenz- und Intensitätsmodulation ermöglicht (Deoni et al., 2011). So konnten Wermke und Kollegen eindrücklich zeigen, dass der Grundfrequenzverlauf im Säuglingsschrei bereits ab der 9. Lebenswoche stärker als zuvor moduliert wird. Zusätzlich kommt es zu einer Entkopplung von der zugehörigen Intensitätskontur, was ebenfalls durch zunehmende neurophysiologische Kontrollmechanismen erklärt wird (Wermke et al., 2002).

Die höchste Ebene in der Hierarchie des Stimmkontrollsystems ist durch den LMC und seine Input- und Outputstrukturen repräsentiert. Anfänglich nur in Tiermodellen als „Vokalisationsverhalten“ im inferioren Teil des motorischen Cortex unterhalb der Repräsentation des Kiefers und oberhalb der Repräsentation der Lippenmuskulatur beschrieben, wurde die laryngeale Region mittlerweile im primären motorischen Cortex, vorherrschend im Areal 4p, lokalisiert (Loucks T. M., Poletto C. J., Simonyan K., Reynolds C. L., & Ludlow C. L., 2007).

Funktionell interagiert der LMC mit einer Vielzahl von Strukturen des benachbarten somatosensorischen Cortex und des inferioren Parietalcortex. Diese sind ausschlaggebend für die Integration propriozeptiven und taktilen Feedbacks aus den orofazialen, respiratorischen und laryngealen Regionen während der Stimmproduktion.

⁵ Beispielsweise kann der Säuglingsschrei auch unabhängig von Schmerz und Hunger produziert werden.

Die Myelinisierung dieser Regionen findet zwischen dem 4. und 6. Monat statt (Deoni et al., 2011). Neuroimaging-Studien weisen darauf hin, dass die ventrale sensorisch-motorische Region nicht nur ein Zusammenspiel motorischer und sensorischer Komponenten des Stimm-Kontroll-Systems (z.B. laryngeal, orofazial, respiratorisch) bereitstellt, sondern auch als ein höheres Zentrum für die Koordination des expiratorischen Luftstroms, welcher die Stimm Lippen in Vibration versetzt, gilt sowie für die Modulation der orofazialen Artikulatoren für die willkürliche Produktion verschiedener Vokalisationstypen zuständig ist (z.B. Bohland & Günther, 2006; Loucks T. M. et al., 2007). Die Myelinisierung innerhalb dieser Strukturen erfolgt zwischen dem vierten und sechsten Lebensmonat und bildet möglicherweise die Grundlage für die Zunahme komplexerer Vokalisationen (marginales Babbeln) während dieses Zeitraums (Deoni et al., 2011).

Der inferiore Parietalcortex – supplementär-motorisches Areal und Gyrus angularis inbegriffen – repräsentiert ein höher gelegenes sensomotorisches Organisationsgebiet für die Koordination von Sprachproduktion und Sprachverständnis. Beispielsweise ist er in komplexe phonologische und semantische Verarbeitungsprozesse involviert (Zheng, Munhall K. G., & Johnsrudel, 2010). Auch diese Areale werden bereits zwischen dem 4. und 6. Monat myelinisiert (Deoni et al., 2011).

Für die Stimm- und Sprachproduktion benötigt der LMC ebenso Informationen aus dem inferioren Gyrus frontalis (IFG). Auch hier laufen Kontrollmechanismen für Sprachverständnis und Sprachproduktion, insbesondere für die hierarchische Gliederung linguistischer und non-linguistischer Sequenzen sowie für die artikulatorische Planung von Sprechbewegungen zusammen (Zheng et al., 2010, Simonyan & Horwitz, 2011). Die enge strukturelle und funktionelle Verbindung zwischen IFG und LMC scheint für die motorische Planung und Verarbeitung aller Komponenten der Sprachproduktion verantwortlich zu sein (Simonyan & Horwitz, 2011). Die Myelinisierung der Strukturen des Frontallappens findet erst zwischen dem 6. und 8. Monat statt und könnte somit zeitlich in engem Zusammenhang mit dem Auftreten erster kanonischer Silbensequenzen stehen, deren Produktion eine gezielte artikulatorische Planung erfordert (Deoni et al., 2011). Diese Interpretation findet sich auch bei Simonyan: „Because the IFG is prominent in the processing of information necessary for long-term preparation of learned oro-motor sequences, it is not surprising that the IFG is usually active during production of long sequences of syllables and only rarely during production of single syllables” (Simonyan & Horwitz, 2011, S. 7).

In neueren Neuro-Imaging-Studien wurden mögliche Unterschiede in intrahemisphärischen Verbindungen (dorsale und ventrale Leitungsbahnen) in fronto-temporalen sprachrelevanten Regionen zwischen Neugeborenen und Erwachsenen untersucht (Brauer, Anwander, Perani, & Friederici A. D., 2013; Dubois, Hertz-Pannier, Dehaene-Lambertz, Cointepas, & Le Bihan, 2006; Leroy et al., 2011; Perani et al., 2011). Aufgrund von Perzeptionsstudien ging man davon aus, dass die sprachrelevanten frontalen und temporalen Regionen im Alter von drei Monaten ähnlich wie beim Erwachsenen durch zwei Faserbündel verbunden sind: über den Tractus arcuatus und den Tractus uncinatus (Dubois et al., 2006). Es wurde gezeigt, dass sich die myelinisierten Fasern des Tractus arcuatus, die den temporalen Cortex mit dem Broca-Areal verbinden, in den ersten Lebensmonaten nicht bis in den IFG (Broca-Areal) erstrecken, sondern nur bis in den prämotorischen Cortex reichen (Brauer et al., 2013; Dubois et al., 2006; Leroy et al., 2011; Perani et al., 2011). Die Verbindung zwischen temporalem Cortex und prämotorischen Cortex erlaubt ein sensorisch-motorisches Mapping, welches äußerst wichtig für die frühe Sprachentwicklung ist. Sie garantiert das sensorisch-motorische Feedback, welches während der vorsprachlichen Entwicklung, insbesondere während der kanonischen Babbelphase, eine herausragende Rolle spielt (Perani et al., 2011).

Die reziproken Verbindungen zwischen LMC und SMA werden auch für die stimm-motorische Planungsausführung benötigt. So konnte z.B. gezeigt werden, dass das supplementär-motorische Areal (SMA) während der Ausführung laryngealer Aufgaben aktiv ist (z.B. Loucks T. M. et al., 2007). Ebenso wurde das SMA als eine motorische Komponente des Sprachüberwachungsnetzwerkes (*speech monitoring network*) identifiziert (van de Ven, Esposito, & Christoffels, 2009). Funktionelle Beziehungen zwischen LMC und SMA sind bereits sehr früh vorrangig links-lateral zu finden und sind – verglichen mit angeborenem laryngealen Verhalten (z.B. der Atmung) – während der Ausführung erlernter Lautproduktionen besonders ausgeprägt: „It, thus, appears that the connectivity between the LMC and the pre-SMA, a region responsible for processing of higher order motor plans for subsequently ordered movement execution, plays a central role in sequencing and initiation of complex learned vocal movements during speech production“ (Simonyan & Horwitz, 2011, S. 7).

Zusätzlich zu den direkten Verbindungen zwischen LMC und laryngealen Motoneuronen des Hirnstamms, ist der LMC Bestandteil eines weit ausgedehnten Netzwerkes subkortikaler Verbindungen, die bis zu den laryngealen Motoneuronen des Hirnstamms

projizieren. Der Putamen (Basalganglien, graue Substanz) erhält die stärkste subkortikale Projektion des LMCs. Da Läsionen des Putamen zu dysphonischen oder dysarthrischen Störungen beim Menschen, nicht aber bei Affen führen, geht man davon aus, dass der Putamen nur bei der Produktion erlernter (spontaner) Vokalisationen involviert ist (Jürgens, 2002).

Die direkten Verbindungen des LMCs mit fast allen funktionellen Substrukturen des Striatums, wie z.B. sensomotorische, assoziative und limbische Regelkreise, repräsentieren wichtige Faktoren der integrativen Kontrolle verschiedener Aspekte der Sprachproduktion. Dazu zählen v.a. motorische Kontrollmechanismen, sowie an der Sprachverarbeitung beteiligte Prozesse der Motivation und Kognition (Simonyan & Horwitz, 2011).

Die Myelinisierung dieser Regionen beginnt um den achten Lebensmonat herum (Deoni et al., 2011). Daraus lässt sich ableiten, dass eine fortgeschrittene Reife und intentionale Funktionalität möglicherweise erst danach besteht und diese dann in direktem Zusammenhang mit der Produktion komplexerer Vokalisationen (kanonisches Babbeln, erste Wörter) sowie der Erweiterung des rezeptiven Wortschatzes in diesem Alter steht.

Eine weitere subkortikale Region, welche direkt mit dem laryngeal-motorischen Cortex verbunden ist, stellt der Thalamus dar. Neben der direkten Funktion der sensomotorischen Stimmkontrolle während erlernter Vokalisation stellt der Thalamus eine wichtige Sendestation für die Integration sensomotorischer Informationen aus den Basalganglien und dem Cerebellum dar. Beide Pfade verbinden den LMC letztlich mit den laryngealen Motoneuronen (Simonyan & Horwitz, 2011).

Zusammengefasst ist die zentrale Kontrolle der menschlichen Stimm- und Sprachproduktion hierarchisch organisiert. Die höchste Ebene dieses Kontrollsystems stellt der LMC dar. Er ist unverzichtbar für die Kontrolle erlernter, jedoch nicht angeborener Vokalisationen und hat sowohl direkte Verbindungen zu den laryngealen Motoneuronen des Hirnstamms als auch zu einem weitläufigen Netzwerk kortikaler und subkortikaler Verbindungen, welche den LMC indirekt mit den laryngealen Motoneuronen verbindet. Somit wird die willkürliche Stimmproduktion durch den LMC kontrolliert und über vielzählige Pfade bis zu den laryngealen Motoneuronen des Hirnstamms ausgeführt (Simonyan & Horwitz, 2011). Wie und wann genau sich dieses Netzwerk ausbildet ist noch weitgehend unerforscht, dies erfolgt jedoch mit hoher Wahrscheinlichkeit in großen Teilen bereits in dem in der vorliegenden Arbeit untersuchten Entwicklungszeitraum.

2. Die Entwicklung produktiver Fähigkeiten zwischen dem 3. und 7. Monat

Ziel dieses Kapitels ist es, die Entwicklung der produktiven vorsprachlichen Fähigkeiten im Alter zwischen 3 und 7 Monaten auf verschiedenen Ebenen darzustellen und eine Zusammenfassung der Ergebnisse bisher durchgeführter Studien zu geben. Dabei werden die frühkindlichen Vokalisationen zunächst chronologisch nach Alter auf Grundlage phasenbasierter Erwerbsmodelle mit entsprechender Beschreibung des Lautrepertoires eingeteilt und klassifiziert (siehe Kapitel 2.1). Im Anschluss daran folgt ein Literaturüberblick zur Entwicklung der akustischen und artikulatorischen Eigenschaften in Komfortvokalisationen im Alter von 3-7 Monaten (siehe Kapitel 2.2 und 2.3). Diese Befunde dienen der vorliegenden Studie als Vergleichsgrundlage und lieferten einen wichtigen Beitrag zur Interpretation der Ergebnisse der vorliegenden Analyse.

2.1 Kategorisierung der Lautproduktionen auf Grundlage phasenbasierter Modelle

Im zweiten Lebensmonat beginnen Säuglinge bereits aktiv in vokale Interaktion mit ihrer sozialen Umgebung zu treten (Papousek, 2014). In diesem Kapitel sollen bestimmte Entwicklungsabschnitte der produktiven vorsprachlichen Fähigkeiten zwischen dem 3. und 7. Lebensmonat charakterisiert werden.

Wie in Kapitel 1 bereits erwähnt wurde, herrscht ein allgemeiner Konsens darüber, dass sich zwischen dem dritten und siebten Lebensmonat eine Reihe von anatomischen und neurophysiologischen Veränderungen vollziehen, welche die Entwicklung vorsprachlicher Lautproduktionen beeinflussen (Masataka, 2003).

Diese morphologischen Entwicklungen werden in Zusammenhang mit bestimmten Phasen der frühen sprachlichen Entwicklung gesehen: In den letzten ca. 40 Jahren wurden die Beschreibung und die Darstellung des Ablaufs dieser Phasen zum Forschungsgegenstand verschiedener Autoren. In diesem Kapitel werden hauptsächlich drei in der vorhandenen Literatur sehr häufig zitierte Modelle präsentiert (Koopmans-van Beinum & van der Stelt, 1986; Oller, 1980; Stark, 1980). An mancher Stelle werden diese

2. Entwicklung produktiver Fähigkeiten zwischen dem 3. und 7. Monat

durch neuere Untersuchungen ergänzt. Der Hauptfokus liegt dabei auf den Phasen, die zeitlich in den Untersuchungszeitraum der vorliegenden Arbeit (3.-7. Monat) fallen. Die Phasen vorher und nachher werden aus Gründen der Vollständigkeit erwähnt, aber inhaltlich nicht weiter vertieft.

Die Modelle der frühen sprachlichen Entwicklung beschreiben bis zu sechs aufeinanderfolgende Erwerbsphasen, welche jeweils mit dem erstmaligen Auftreten neuer Vokalisationstypen beginnen. Des Weiteren erfassen sie das Alter, in dem bestimmte Vokalisationstypen erstmalig auftreten (Nathani, Ertmer, & Stark, 2006). Dabei unterscheiden sich die Modelle teilweise in der Phaseneinteilung, in den Klassifikationskriterien und in der verwendeten Terminologie. Im Wesentlichen lassen sich ein (meta-)phonologisch orientierter Klassifikationsansatz (Oller 1980, Stark 1980) und ein sensorisch-motorischer Klassifikationsansatz (Koopmans van Beinum & Van der Stelt 1986) voneinander unterscheiden (Koopmans-van Beinum, Clement, & Van den Dikkenberg-Pot, 2001).

Beim sensorisch-motorischen Ansatz (Koopmans van Beinum & van der Stelt 1986) bildet die phonatorische und/oder artikulatorische Aktivität während der Lautproduktion die Grundlage für die Klassifizierung der vorsprachlichen Vokalisationen. Dabei wird von jeglicher Interpretation der Qualität der daraus resultierenden Lautproduktion abgesehen. So wird beispielsweise nicht zwischen konsonantischen und vokalischen Elementen unterschieden (Koopmans-van Beinum & van der Stelt, 1986).

Im (meta-)phonologischen Ansatz (Oller 1980, Stark 1980) hingegen ist v.a. bei der Klassifikation der Vokalisationen in marginale, kanonische und bunte Babbelvokalisationen die Ausführungsqualität der Lautproduktion von Bedeutung. So werden zur Einordnung der Lautproduktionen definierte zeitliche und segmentale Kriterien – in Anlehnung an die natürlichen Sprachen – aufgeführt.

Trotz der Unterschiede beanspruchen beide Ansätze für sich, dass sie die Lautproduktionen aus der Perspektive des kindlichen Erwerbssystems heraus interpretieren und klassifizieren: "[...] focused upon the manner in which the infant's own system develops, not upon his acquisition of an adult system" (Stark, 1980). Sie wollen den frühen Spracherwerb aus der Rolle des Kindes charakterisieren, das seinen Output entsprechend seinen Möglichkeiten modifiziert, indem es zunehmend mehr Kontrolle über seine Stimmlippenschwingung, die Bewegungen seiner supraglottalen Artikulatoren sowie der Fähigkeit diese beiden Systeme miteinander zu koordinieren erlangt.

2. Entwicklung produktiver Fähigkeiten zwischen dem 3. und 7. Monat

Das Modell von Stark (1980) basiert auf sechs Phasen und entstand auf der Grundlage einer Längsschnittstudie zu auditiven und akustischen Vokalisationseigenschaften bei allerdings nur zwei weiblichen englisch-sprachigen Probanden im ersten Lebensjahr.

Phase I (*Reflexive vocalization*) erstreckt sich von der Geburt bis zum Alter von 1,5 Monaten und umfasst v.a. Schrei-Laute, Diskomfort-Laute und vegetative Laute. Diese Laute werden hauptsächlich als vokalisches beschrieben. Vegetative Laute enthalten aber bereits auch konsonanten-ähnliche Elemente. Diese werden im Allgemeinen als „stops“, „clicks“, „friction noises“⁶ und „trills“⁷ (stimmlos) beschrieben. Nasale und „liquids“ (stimmhaft) werden in diesem Alter eher selten beobachtet (Stark, 1978; Stark, 1980). Die konsonanten-ähnlichen Elemente treten meistens am Beginn oder Ende des Lautes auf, wobei die „friction noises“ und „trills“ in der Regel eine Länge von 100 ms aufweisen.

In Phase II (*Gurren und Lachen*) – im Alter von 1,5 bis 3 Monaten – treten typischerweise erstmals Komfortvokalisationen und Lachen auf. Diese werden zunächst in Interaktion mit einem Erwachsenen oder anderem Kind produziert, später aber auch in Situationen, die keine soziale Interaktion beinhalten. Die Komfortvokalisationen in Phase II beinhalten die gleichen konsonantenähnlichen Elemente wie in Phase I, allerdings werden diese zunehmend auch stimmhaft und neben Anfangs- und Endposition auch in medialer Position des Lautes produziert. [...] It was found that the consonantal noises that were present only in vegetative sounds in the first 6 weeks of life were incorporated into comfort sounds produced after 6-8 weeks of age (Stark, 1980, p. 79).

Phase III (*Vocal play*) zwischen dem 4. und 7. Monat wird bei Stark (1980) als „spielerisches Vokalisationsverhalten“, welches „squealing“, „growling“, „yelling“, die Produktion von Enge-Lauten⁸ im Mundraum oder Pharynx sowie „nasal murmurs“ enthalten kann⁹. Bei Oller (1980) wird diese Phase auch als marginales Babbeln beschrieben. Dabei treten innerhalb eines Lautes sowohl vokalische als auch konsonantische Elemente auf, allerdings erfüllen diese Silben in ihren akustischen Eigenschaften noch nicht die Kriterien von Silben in der Erwachsenensprache.

⁶ „Friction noises“ (Reibelaute) werden durch eine Verengung im Vokaltrakt oberhalb der Glottis gebildet (Stark, 1980).

⁷ „Trills“ (Zungenspitzenlaute) werden ähnlich wie Reibelaute gebildet, allerdings wird die Verengung mehrere Male in Folge geöffnet, sodass die Zungenspitze mehrere Male den Gaumen oder die Lippen berührt (Stark, 1980).

⁸ Dabei werden oft Geräusche erzeugt, die durch ein Blasen von Luft oder Speichel durch die Engstelle hervorgerufen werden (Stark, 1980).

⁹ Zur Erläuterung der Vokalisationstypen siehe Fußnoten 11-15 auf Seite 25.

2. Entwicklung produktiver Fähigkeiten zwischen dem 3. und 7. Monat

Verglichen mit Vokalisationen aus Phase II sind die konsonanten-ähnlichen Elemente in den Komfortvokalisationen aus Phase III wesentlich länger (durchschnittlich mehr als 10-mal so lang) und umfassen teilweise bis zu 1,5 Sekunden. Häufig treten Frikative, Affrikate und Schnalzlaut auf. Sie weisen eine Konzentration der Energie in den höheren Frequenzen auf.

In Phase IV (*Reduplicated babbling*) zwischen dem 6. und 10. Monat treten zum ersten Mal kanonische CV-Silben (z.B. ba-ba-ba, da-da-da) auf. „In reduplicated babbling, sufficient control over phonation is needed to coordinate it with repeated opening and closing of the vocal tract above the glottis and to impose at least a simple fundamental frequency contour upon the syllable series produced“ (Stark, 1980, p. 81). Durch ihre zeitlichen Eigenschaften (C-V-Übergang) werden kanonische Silben bereits als erwachsenensprachlich wahrgenommen.

Phase V (*Nonreduplicated babbling*) zwischen 10 und 14 Monaten ist durch die Produktion verschiedener Konsonanten und Vokale innerhalb einer Äußerung gekennzeichnet.

In Phase VI (*Single-word productions*) werden ab dem 10. Monat bereits erste Wörter (Protowörter) produziert. Sie haben Symbol-Charakter und werden wiederkehrend für gleiche Objekte und Situationen verwendet (Stark, 1980).

Ollers (1980) Untersuchungen basieren größten Teils auf Beobachtungen seiner eigenen Kinder (englisch-sprachig). Die Phasen, die er beschreibt, überschneiden sich teilweise mit denen von Stark (1980).

In Phase I (*Phonation Stage*) zwischen Geburt und erstem Monat werden laut Oller (1980) neben reflexiven Lauten und Diskomfort-Lauten bereits auch „sprach-ähnliche“ Laute produziert. Er nennt diese Laute „Quasi-Vokale“ („quasi-resonant nuclei“). „Quasi-resonant nuclei include normal phonation (not vocal fry, not breathy voice, etc.), but do not seem to involve any systematic contrast between opening and closure of the vocal tract, and do not make use of the full potential of the vocal cavity to function as a resonating tube“ (Oller, 1980, p. 95). Die akustische und phonetische Qualität der Quasi-Vokale ergibt sich daraus, dass sie mit geschlossenem oder fast geschlossenem Mund produziert werden. Bis zum zweiten Lebensmonat können vereinzelt auch bereits „voll-resonante“ Vokale („fully resonant nuclei“) produziert werden. Deren Auftreten beschreibt Oller (1980) jedoch als zufällig und willkürlich.

Zwischen dem zweiten und dritten Lebensmonat (*Goo-Stage*) treten erstmals sogenannte Gurr-Laute auf (Phase II). Sie bestehen aus postalveolaren (meist velaren) konsonanten-ähnlichen Elementen (oft stimmhafte Frikative¹⁰) in Kombination mit Quasi-Vokalen. Auch wenn die Gurr-Laute bereits eine primitive Silbenstruktur aufweisen, besitzen sie noch keine zeitlich regulären Mechanismen bezüglich der Öffnung (QV) und Schließung (konsonanten-ähnliche Elemente) wie es in der Erwachsenensprache der Fall ist. „The temporal limitations of Goos should be kept in mind to avoid the spurious conclusion that infants at this stage have already achieved a full phonetic competence“ (Oller, 1980, p. 96).

Um den fünften Monat (4-6 Monate) erweitert sich das Vokalisationsrepertoire enorm (Phase III, *Expansion Stage*). Charakteristisch für diese Phase ist die Produktion von vokal-ähnlichen (voll-resonanten¹¹) Elementen, die Produktion von Raspberry-Vokalisationen¹², Squealing¹³, Growling¹⁴, Yelling¹⁵ und marginales Babbeln. Letzteres beschreibt Oller (1980) als eine Abfolge von konsonanten-ähnlichen Elementen und voll-resonanten vokal-ähnlichen Elementen. Dabei merkt auch er an, dass die zeitlichen Eigenschaften der marginalen Silbe nicht den zeitlichen Eigenschaften in späteren Phasen oder der Erwachsenensprache entsprechen. „Impressionistically, marginal babbling includes slow or shaky transitions between consonant-like and vowel-like elements“ (Oller, 1980, p. 98). Im Vergleich zu den anderen Vokalisationstypen dieser Phase tritt das marginale Babbeln verhältnismäßig selten auf. Zwischen sieben und 10 Monaten (*Canonical Stage*, Phase IV) beginnt das Kind erstmals kanonische Silben zu bilden, die aus konsonantischen und vokalischen Elementen bestehen. Die kanonische Silbe erfüllt bereits die zeitlichen Vorgaben der Erwachsenensprache. Oller legt einige Kriterien fest, welche eine kanonische Silbe erfüllen muss (Oller, 1986):

- Ein Intensitätsabfall von mindestens 10 dB zwischen konsonantischem und vokalischem Element
- Silbenlänge zwischen 100 und 500 ms

¹⁰ Vergleiche „friction noises“ bei Stark Stark (1980).

¹¹ Fully resonant Nuclei (Oller, 1980): vowel-like elements that systematically include strong resonances above 1200 Hz Oller (1980).

¹² Raspberry vocalizations (Oller, 1980): consisting of bilabial or labiodental trills or vibrants (referred to as „trills“ or „friction noises“, cf. Stark, 1980).

¹³ Squealing utterances (Oller, 1980): often begins or ends in a normal vocal register but, at some point, enter falsetto or a highly tense maximal pitch register

¹⁴ Growling (Oller, 1980): seems to be the inverse of Squealing; consists of very low pitch, often creaky-voice vocalizations.

¹⁵ Yelling (Oller, 1980): consists of high amplitude nondistress vocalization.

2. Entwicklung produktiver Fähigkeiten zwischen dem 3. und 7. Monat

- Ein geöffneter Vokaltrakt während der Produktion der Nuclei (voll-resonant)
- Mindestens ein konsonantisches Element pro Silbe
- Ein Formantübergang von weniger als 120 ms
- Der Intensitätshub innerhalb einer Silbe darf 30 dB nicht überschreiten
- Der F_0 -Range darf 2 Ganztöne nicht überschreiten

In der kanonischen Phase werden oft Silben mit gleichbleibenden konsonantischen und vokalischen Elementen produziert (redupliziertes Babbeln). Zwischen 11 und 12 Monaten (*Variegated Babbling Stage*), kurz bevor erste Wörter auftreten, produziert das Kind dann zunehmend Silben mit unterschiedlichen konsonantischen und vokalischen Elementen (Oller, 1980).

Koopmans-van Beinum und Van der Stelt (1986) präsentieren ein Modell, welches auf einer Längsschnittstudie an ebenfalls nur zwei holländischen männlichen Probanden im ersten Lebensjahr basiert. Dabei wurden sowohl phonatorische als auch artikulatorische Bewegungen innerhalb einer Expirationsphase festgehalten und beschrieben und die Annahme einer kontinuierlichen und parallelen Entwicklung auf laryngealer und supralaryngealer Ebene bestätigt. „Comparing and combining the results on both infants revealed that a phonatory and articulatory description of sounds produced in the first months shows a systematic continuity“ (Koopmans-van Beinum & van der Stelt, 1986, p. 39).

Koopmans-van Beinum und Van der Stelt (1986) haben drei Meilensteine der frühen Vokalisationsentwicklung festgelegt:

- Produktion mehrerer „glottal stops“ innerhalb einer Expirationsphase
- Beginn der Produktion artikulatorischer Elemente
- Produktion sich wiederholender artikulatorischer Elemente

Basierend auf dem Quelle-Filter-Modell nach Fant (1960) wurde zusätzlich ein Transkriptionssystem entworfen. Als Quelle werden dabei die Lunge und der Larynx (laryngeales System) betrachtet, da beide Organe für die Phonation ausschlaggebend sind. Den Filter bildet der Vokaltrakt (supralaryngeales System), da er für die artikulatorische Konfiguration des Phonationsstroms verantwortlich ist (Titze, 1994).

In Phase I (0-1,5 Monate) werden hauptsächlich Laute produziert, in denen der Phonationsstrom nicht durch artikulatorische Elemente unterbrochen wird.

2. Entwicklung produktiver Fähigkeiten zwischen dem 3. und 7. Monat

Phase II (1,5-2,5 Monate) ist durch das vermehrte Auftreten von Unterbrechungen im Phonationsstrom innerhalb einer Expirationsphase gekennzeichnet (*Meilenstein 1*). Dabei handelt es sich jedoch noch nicht um artikulatorische Elemente.

Diese treten erstmals in Phase III (2,5-4,5 Monate) auf. In diesem Zeitraum wird ein artikulatorisches Element mit einer unterbrochenen oder durchgängigen Phonation innerhalb einer Expirationsphase verbunden (*Meilenstein 2*).

In Phase IV (4,5-6 Monate) kommt es zu einem enormen Anstieg der Auftrittshäufigkeit artikulatorischer Elemente und zu einer Abnahme von Äußerungen mit durchgängiger Phonation ohne Artikulation. „However, this period is characterized by all kinds of variations in the phonatory domain concerning intonation, duration and intensity“ (Koopmans-van Beinum & van der Stelt, 1986, p. 43).

Phase V (7-12 Monate) ist durch eine reduplizierte Anordnung artikulatorischer Elemente innerhalb einer Expirationsphase gekennzeichnet. Auch hier können diese mit unterbrochener oder durchgängiger Phonation kombiniert werden (*Meilenstein 3*).

In Phase VI werden erstmals bedeutungstragende Wörter produziert (Koopmans-van Beinum & van der Stelt, 1986).

Zusammengefasst beschreiben die drei Entwicklungsmodelle bis zu sechs regelhafte Erwerbsphasen, während derer der Säugling in Abhängigkeit von der Kontrolle seiner Stimmlippenschwingung, der Bewegungsfähigkeit seiner supralaryngealen Artikulatoren sowie der Fähigkeit diese beiden Systeme miteinander zu koordinieren seine Vokalisationen manipuliert. Sie erlauben es das frühkindliche Lautrepertoire anhand geeigneter Unterscheidungsmerkmale zu kategorisieren und einem Reifestadium zuzuordnen. Trotz spezifischer Unterschiede der Modelle, sind sich die Autoren einig, dass es sich um typische Entwicklungsphasen handelt, welche im Sinne eines angeborenen Entwicklungsprogramms universell von allen Kindern durchlaufen werden. Dabei herrschen in den einzelnen Entwicklungsphasen zwar bestimmte Vokalisationstypen vor, jedoch werden auch (noch lange) typische Lautäußerungen aus früheren Entwicklungsphasen parallel produziert (Papousek, 1994; Stark, 1980).

Während in der ersten Phase sog. *reflexive Laute* und *rein vokalische Vokalisationen*, bei denen der Phonationsstrom nicht unterbrochen wird, dominieren, treten in der zweiten Phase bereits sog. *Gurrlaute mit primitiver Silbenstruktur* und ersten konsonantischen Elementen auf, welche im hinteren Bereich des Vokaltraktes gebildet werden. In der dritten Phase erweitert sich das Vokalisationsrepertoire enorm. Charakteristisch ist die Produktion von sog. *squeals, raspberries und marginalen*

Babbel-Lauten. In den darauffolgenden Phasen kommt es zu einem enormen Anstieg in der Produktion artikulatorischer Elemente. Typisch ist das Auftreten *kanonischer Silben* (ba-ba-ba) gefolgt von *ersten Wörtern* (Koopmans-van Beinum & van der Stelt, 1986; Oller, 1980; Stark, 1980).

2.2 Akustische Eigenschaften von Komfortvokalisationen zwischen dem 3. und 7. Monat

Im letzten Abschnitt wurde deutlich, dass sich die neurophysiologischen und morphologischen Veränderungen des Phonations- und Artikulationssystems auditiv-perzeptiv wahrnehmbar im Lautrepertoire widerspiegeln und beispielsweise transkribiert werden können. Mittels Spektral- und Melodieanalyse ist es möglich diese Entwicklung in Form von Veränderungen akustischer Lautparameter auch objektiv zu charakterisieren.

Eine Schwierigkeit dabei ist der Umstand, dass sich die bisher durchgeführten Studien zur Untersuchung akustischer Kenngrößen frühkindlicher Vokalisationen (z.B. mittlere F_0 , F_0 -Range, Melodiestructur, mittlere Vokalisationslänge) in ihrer Methodik teilweise deutlich voneinander unterscheiden. Auch limitieren methodische Schwächen ihre Aussagekraft: Beispielsweise wurden unterschiedliche Vokalisationstypen (Cry, Komfort) in Untersuchungen einbezogen und nicht differenziert analysiert. Einige Autoren konnten signifikante Unterschiede der F_0 -Daten in Abhängigkeit vom Vokalisationstyp beobachten. Mittlere F_0 und F_0 -Range liegen z.B. in Schrei-Vokalisationen höher als in Nicht-Schrei-Vokalisationen (Fuller & Horii, 1988; Murry, Hoit-Dalgaard, & Gracco, 1983; Rothgänger, 2003). Auch die Analyseverfahren unterscheiden sich. So verwendeten einige Autoren rein auditiv-perzeptive Verfahren, während andere Autoren auf spektralanalytische Verfahren zurückgriffen. Dabei ist zu bedenken, dass die rein auditiv-perzeptiven Verfahren im Säuglingsalter stark durch die subjektive Wahrnehmung des Untersuchers beeinflusst sein können und es somit zu psychoakustisch bedingten Fehlinterpretationen kommen kann. Ebenso wurden die Analyseeinheiten von Autoren unterschiedlich definiert. Einige definierten ihre Analyseeinheiten mittels festgelegter Pausen zwischen den einzelnen Vokalisationen in den Aufnahmesequenzen, andere hingegen definierten sie als Äußerung innerhalb einer Expirationsphase. Das kann zu erheblichen Unterschieden in der Bezugsgröße für akustische Kenngrößen der vorsprachlichen Vokalisationen führen. Wichtig erschien es auch bei der hier beschriebenen Zusammenfassung bisheriger Messungen zu beachten,

dass die Untersuchungszeiträume der einzelnen Studien sehr variabel waren. Die Methodik und zu beachtende Besonderheiten (Probandenzahl, Definition von Analyseeinheiten usw.) der einzelnen Studien können Tabelle 30 (siehe Anhang) entnommen werden. Die Literaturrecherche ergab, dass sich die anatomischen und neurophysiologischen Veränderungen im ersten Lebensjahr unterschiedlich stark auf die verschiedenen akustischen Kenngrößen auswirken.

Studien, die die mittlere F_0 untersuchten, führen zusammengefasst zu dem Ergebnis, dass die mittlere Grundfrequenz¹⁶ von Komfortvokalisationen im ersten Lebensjahr etwa zwischen 300 Hz und 450 Hz liegt. Die meisten Autoren stellten keine signifikanten Veränderungen in diesem Zeitraum fest (Delack, 1976; Delack & Fowlow, 1978; Iyer & Oller, 2008; Kent & Murray, 1982; Laufer & Horii, 1977; Murry et al., 1983; Rothgänger, 2003; Ruzza, Rocca, Lenti Boero, & Lenti, 2003; Scheiner, Hammerschmidt, Jürgens, & Zwirner, 2002; Wermke et al., 2013). Lediglich eine Studie konnte einen signifikant steigenden Trend aufdecken (Clement, 2004).

Dies spricht dafür, dass die laryngealen Regelmechanismen bereits zu einem relativ frühen Zeitpunkt reif sind. Gruppenunterschiede bei Kindern mit unterschiedlichem muttersprachlichem Input sind bisher nicht beschrieben (Ruzza et al., 2003; Wermke et al., 2013).

Allerdings wurden signifikante Geschlechtsunterschiede für die mittlere F_0 gefunden. In der Längsschnittuntersuchung von Delack und Fowlow (1976) verhielt sich die mittlere F_0 im ersten Lebensjahr relativ stabil bei 355 Hz, jedoch lag die F_0 der weiblichen Probanden 20-25 Hz höher als die der männlichen Probanden. Im 6. Monat fand sich bei beiden Geschlechtern ein F_0 -Maximum.

Clement (2004) konnte als einzige Autorin in ihren Längsschnittdaten (2-12 Monate) einen Alterseffekt bezüglich der mittleren F_0 aufdecken. Dabei lagen die Werte der mittleren F_0 im zweiten Monat signifikant unter den Werten zu späteren Zeitpunkten.

Die Auswertung der Fachliteratur ergab ein unzureichendes Bild von Veränderungen der mittleren F_0 im ersten Lebensjahr. Es gibt zu wenig systematische und ausreichend genaue Untersuchungen. Da die F_0 eine sehr wesentliche Messgröße für die Charakterisierung der laryngealen Regelmechanismen ist, sind aus entwicklungsdiagnostischer Sicht weitere spezifische und signalanalytisch zuverlässigere Analysen notwendig. Die vorliegende Studie liefert dazu einen wichtigen Beitrag.

¹⁶ Im Folgenden abgekürzt durch „mittlere F_0 “.

2. Entwicklung produktiver Fähigkeiten zwischen dem 3. und 7. Monat

Der Variationsbereich der F_0 innerhalb einer Äußerung¹⁷ scheint vom Alter abhängig zu sein. Er steigt kontinuierlich bis zum ersten Lebensjahr an, wobei es zwischen dem 4. und 6. Lebensmonat zu besonders hohen Werten kommt, die typisch für die Phase des *Expansion Stages* (Oller, 2000) und die während dieser Zeit vorherrschend produzierten Vokalisationstypen sind (Clement, 2004; Delack & Fowlow, 1978; Laufer & Horii, 1977; Murry et al., 1983). Während dieser Phase kommt es zu einer spielerischen Erkundung des Resonanzraums und des F_0 -Ranges und Vokalisationstypen, die auf Stimmregisterwechsel basieren, wie *squealing*, *growling*, *yelling* und *raspberries*, treten erstmals auf. Dies spricht dafür, dass sich der Variationsbereich der laryngealen Aktivität durch eine zunehmende neurophysiologische Kontrolle im Verlauf dieser Phase vergrößert und zeitgleich zu einer Erweiterung des Vokalisationsrepertoires führt. Ab dem 6. Lebensmonat wurden erste geschlechtsspezifische Unterschiede gefunden. So zeigen Delack und Follow (1978), dass der F_0 -Range der weiblichen Probanden in der zweiten Hälfte des ersten Lebensjahres weiterhin ansteigt, während er bei den männlichen Probanden abfällt (Delack & Fowlow, 1978).

Auch die Zeitfunktion der F_0 , die Melodie, wurde als wichtiges prosodisches Element bereits in einigen Studien untersucht. So konnte anhand von Schrei-Vokalisationen beispielsweise gezeigt werden, dass sich die Melodiestructur, welche sich objektiv durch den Melodieverlauf und die zeitliche Segmentierung innerhalb einer einzelnen Vokalisation beschreiben lässt, in den ersten Lebensmonaten nach einem angeborenen Entwicklungsprogramm von zunächst einfachen (einböigen) hin zu zunehmend komplexeren (mehrböigen) Strukturen entwickelt (Wermke, 1993, 1994; Wermke et al., 2002; Wermke & Mende, 2011).

Die Komplexität dieser Melodiestructuren reflektiert dabei die Modulationskapazität des Säuglings hinsichtlich der Grundfrequenz (F_0). So kommen auch Untersuchungen an Komfortvokalisationen zu dem Ergebnis, dass die Komplexität der Melodiestructur in den ersten sechs Lebensmonaten zunimmt und es eine Kontinuität in der Entwicklung vom Weinen zu Komfortvokalisationen gibt (Clement, 2004; Delack & Fowlow, 1978; Hsu, Fogel, & Cooper, 2000; Kent & Murray, 1982). „Thus, from a strictly acoustic perspective, early cry and the later vocalizations of cooing and babbling appear to be vocal performances in continuity“ (Kent & Murray, 1982, p. 353).

Zusätzlich steht die individuelle Produktion komplexer melodischer Strukturen in Zusammenhang mit der Mutter-Kind-Interaktion. Legerstee et al. (1991) konnten zeigen,

¹⁷ Im Folgenden abgekürzt durch „ F_0 -Range“

2. Entwicklung produktiver Fähigkeiten zwischen dem 3. und 7. Monat

dass signifikant mehr melodische Strukturen produziert wurden, wenn die Mutter aktiv in Interaktion mit ihrem Kind trat. „Although it was difficult to discern audible differences in the vocalizations during the first months of life, by 7 weeks these melodic vocalizations began to be produced primarily in a ‘conversational’ context” (Legerstee, 1991, p. 339).

Eine weitere wichtige akustische Kenngröße ist der Dauer der Vokalisation. Die Komplexität der Melodiestructur innerhalb einer Äußerung hat großen Einfluss auf die mittlere Lautlänge: je komplexer, desto länger (Bloom, 1989). Somit ist es nicht verwunderlich, dass auch die Werte der mittleren Vokalisationslänge im ersten Lebensjahr kontinuierlich um ca. 50% ansteigen. Dies wird durch eine zunehmende Kontrolle der Abstimmung zwischen Stimm- und Atmungsmechanismen möglich.

Die Werte für die mittlere Lautlänge (über das erste Lebensjahr gemittelt) variieren zwischen 550 ms und 999 ms, wobei studienübergreifend Maximalwerte von bis zu 1400 ms im Alter von 3 und 6 Monaten gefunden wurden (Bettany, 2004; Clement, 2004; D’Odorico, Franco, & Vidotto, 1985; Delack & Fowlow, 1978; Kent & Murray, 1982; Laufer & Horii, 1977; Murry et al., 1983; Rothgänger, 2003; Ruzza et al., 2003).

2.3 Artikulatorische Fähigkeiten zwischen dem 3. und 7. Monat

Neben den Modulationsfähigkeiten auf laryngealer Ebene entwickeln sich im ersten Lebensjahr auch die artikulatorischen Fähigkeiten auf supralaryngealer Ebene. Dabei steigt sowohl die Anzahl an artikulierten Vokalisationen als auch die Anzahl verschiedener Konsonanten-Typen an. Dennoch weisen im ersten Lebensjahr nur ca. ein Drittel aller produzierten Äußerungen konsonantische Elemente auf (Clement, 2004; Irwin, 1946; Stoel-Gammon & Otomo, 1986).

Während in den ersten sechs Monaten 60-80% der artikulatorischen Elemente im dorsalen Bereich (uvular, velar, glottal, pharyngeal) des Vokaltraktes gebildet werden, treten diese in den späteren Phasen kaum noch auf (10-20%). In den ersten beiden Lebensmonaten werden sogar 98% der artikulatorischen Elemente im dorsalen Bereich gebildet. Umgekehrt verhält es sich mit der Produktion zentral gebildeter (alveolar, dental, postdental, apikal) konsonantischer Elemente. Diese treten vor dem 6. Monat kaum auf (<10%) und stellen später den am häufigsten (ca. 70%) produzierten Konsonanten-Typ dar. Die konsonantischen Elemente der vorderen Artikulationszone (labial, labio-dental) nehmen eher eine Mittelposition ein. Bis zum sechsten Lebensmonat kommt es zu einem Anstieg in der Produktion labialer Konsonanten. So

2. Entwicklung produktiver Fähigkeiten zwischen dem 3. und 7. Monat

werden im Alter von 6 Monaten fast 50%, mit 7,5 Monaten sogar 67% der Artikulationen im frontalen Bereich gebildet. Dieser Peak überschneidet sich zeitlich eindeutig mit dem Eintritt in die kanonische Babbelphase, wo Äußerungen wie „ba-ba-ba“ – bestehend aus Silbensegmenten – vorherrschend auftreten. Nach dem 9. Monat fällt der prozentuale Anteil labialer Konsonanten auf ca. 10% ab und die frontal gebildeten Artikulationen werden scheinbar durch Artikulationsbewegungen im zentralen Artikulationsbereich abgelöst. Dabei treten initiale Konsonanten häufiger auf als mediale. Die finale Position wird im ersten halben Jahr fast gar nicht belegt. Letztere nimmt jedoch in der zweiten Hälfte des ersten Lebensjahres enorm zu (Clement, 2004; Grenon, Benner, & Esling, 2007; Irwin, 1946, 1947b, 1951; Smith & Oller, 1981; Stoel-Gammon, 1988).

Zusammengefasst werden im ersten Lebenshalbjahr hauptsächlich Konsonanten produziert, die im hinteren Bereich des Vokaltraktes und an initialer Lautposition gebildet werden. Danach kommt es hinsichtlich der Bildung artikulatorischer Elemente zu einer Verschiebung in Richtung vorderer und zentraler Artikulationszonen, welche auch mediale und finale Lautpositionen belegen.

3. Methodologie und Forschungsfragen

Theoretische Überlegungen

Mittlerweile herrscht ein allgemeiner Konsens darüber, dass früh auftretende Sprachstörungen bis in das frühe Erwachsenenalter persistieren und Auswirkungen auf die kognitive, soziale und emotionale Entwicklung des betroffenen Kindes haben können. Obwohl die Spontansprache bis in das Schulalter meist unauffällig wird, zeigen die Kinder oft Probleme bei der Produktion und dem Verständnis komplexer linguistischer Strukturen, v.a. beim Schriftspracherwerb. Bei einer Zurückbildung der Sprachörungssymptomatik bis in das Kindergartenalter gleichen die Entwicklungschancen der betroffenen Kinder jedoch fast denen unauffälliger (Michelsson, Eklund, Leppänen, & Lyytinen, 2002). Daher geht der Trend in der Spracherwerbsforschung immer mehr hin zur Erforschung von Frühindikatoren für spätere Sprachentwicklungsstörungen. Als Kriterium für den Beginn der Sprachentwicklung gilt jedoch v.a. in der klinischen Sprachentwicklungsdiagnostik gewöhnlich immer noch das erste sinnvolle, intentional gebrauchte Wort (Papousek, 1994). Aus den Heilmittelberichten einzelner Krankenkassen der letzten Jahre geht hervor, dass fast ausschließlich logopädische Behandlungen ab einem Alter von 5 Jahren verordnet werden. Sprachtherapie im Alter von unter 3 Jahren wird kaum verordnet (Waltersbacher, 2020). Somit verstreicht die sensible Phase des Spracherwerbs oft weitgehend ungenutzt.

Gründe dafür sind in fehlenden oder mangelnden Methoden zur zuverlässigen Früherkennung und effektiven Frühintervention während dieser Phase des Spracherwerbs zu sehen. Es gibt zwar bereits Eltern-Fragebögen zur evidenzbasierten Diagnose von „Late Talkers“, doch auch diese haben ihre Einschränkungen. So ist beispielsweise der positive Vorhersagewert bei 2-jährigen Kindern für spätere Sprachentwicklungsstörungen durch solche Sprachfrüherkennungsmethoden mit 50% relativ niedrig (Michelsson et al., 2002). Eine ausreichend sichere Erkennung sprachentwicklungsgestörter Kinder gelingt sogar erst am Ende des dritten Lebensjahres (Suchodoletz, 2011). Folglich setzt nur sehr selten eine frühe und gezielte therapeutische Förderung bzw. eine Sensibilisierung des Umfelds durch spezifische Beratung der Eltern oder Elternteraining (z.B. Buschmann, 2017: Heidelberger

Elterntertraining frühe Sprachförderung) oder sprachheilpädagogische Maßnahmen in Betreuungseinrichtungen ein.

Besonders deutlich zeigen sich die Probleme und Möglichkeiten der Frühdiagnostik und Frühförderung am Beispiel von Kindern mit hochgradigen Hörstörungen. Zwar erfolgt die Diagnose und Versorgung von angeborenen Hörstörungen (seit 2009 Neugeborenen-Hörscreening bei U2) in den meisten Fällen bereits innerhalb der ersten 4 bis 8 Monate, dennoch liegen bis zum jetzigen Zeitpunkt keine geeigneten Instrumente oder Verfahren zur Beurteilung produktiver (vor-)sprachlicher Leistungen im ersten Lebensjahr vor, welche beispielsweise die Anpassung entsprechender Hörhilfen bzw. Maßnahmen im Rahmen einer sprachlichen Frühförderung optimieren könnten.

Verschiedene Studien konnten zeigen, dass sich bereits in frühen vorsprachlichen Lautproduktionen Hinweise auf spätere Sprachentwicklungsstörungen bzw. -verzögerungen finden. Wermke und Kollegen konnten anhand von Melodieanalysen in Säuglingsschreien eindrucksvoll belegen, dass die Melodie-Komplexität dieser frühen Schrei-Vokalisationen bereits im 2. Lebensmonat ein Prädiktor für spätere Sprachleistungen sein kann. So haben Kinder, deren Spontan-Schreie (im schmerzfreien Kontext) im 2. Lebensmonat nicht einen bestimmten Melodiekomplexitätsgrad aufweisen, später ein 5-fach höheres Risiko für Sprachentwicklungsstörungen (Wermke et al., 2007). Auch nachfolgende vorsprachliche Phasen liefern Hinweise auf spätere produktive Sprachleistungen. Wichtige Indikatoren scheinen dabei das verzögerte Einsetzen der kanonischen Babbelphase (Oller, Eilers, Neal, & Cobo-Lewis, 1998), (Oller, Eilers, Neal, & Schwartz, 1999) sowie eine geringere phonetische Komplexität der Babbellaute zu sein (Fasolo, Majorano, & D'Odorico, 2008; Stoel-Gammon, 1998). Daher wird der kanonischen Babbelphase in der Fachliteratur auch aus einer frühdiagnostischen Perspektive große Bedeutung beigemessen (Oller et al., 1998). Inwiefern sich auch Komfortvokalisationen früherer vorsprachlicher Stadien als Prädiktoren für den späteren Spracherwerb eignen, ist noch nicht erforscht. Um jedoch zuverlässig anwendbare Verfahren zur Früherkennung von Spracherwerbsstörungen sowie zur Sprachförderung und -therapie entwickeln zu können, ist es unerlässlich, zunächst systematische Untersuchungen dieser Stadien im Rahmen der ungestörten Entwicklung und vor dem Hintergrund verschiedener externer Einflüsse durchzuführen. Nur durch geeignete Längsschnittuntersuchungen kann untersucht werden, welche Mechanismen innerhalb und am Übergang einzelner Phasen (z.B. von der Phase des Säuglingsschreies zur Expansionsphase) greifen und welche potenziellen Zusammenhänge zum späteren Spracherwerb bestehen. Die vorliegende

Arbeit möchte hierzu einen wichtigen Beitrag leisten, indem sie durch die erstmalige systematische und objektive Charakterisierung von Komfortvokalisationen im Alter zwischen drei und sieben Monaten erste objektive Referenzgrößen für Kinder mit deutscher Umgebungssprache liefert.

Zur Erfüllung dieser Zielstellung ergaben sich nach einer ausführlichen Literaturrecherche (Kapitel 2) und darauf basierenden theoretischen Überlegungen mehrere Forschungsfragen. Diese sollen im Folgenden kurz hergeleitet werden. Im Anschluss wird die methodologische Herangehensweise erläutert und reflektiert. Sie soll das Studiendesign, die Probandenauswahl, die Lautauswahl, die Untersuchungsparameter sowie die verwendeten Analyseverfahren begründen.

Ableitung von Forschungsfragen

Die theoretische Grundlage der vorliegenden Arbeit ist die Annahme, dass die vorsprachliche Entwicklung nach einer bestimmten Regelmäßigkeit verläuft, die oft durch verschiedene Phasen beschrieben wird. Die Entwicklungsmuster folgen dabei einem Kontinuitätsprinzip von einfachen hin zu komplexen Strukturen (Koopmans-van Beinum & van der Stelt, 1986; Oller, 1980; Stark, 1980), siehe Kapitel 2. Wermke und Kollegen konnten in einigen Untersuchungen, u.a. zur Melodieentwicklung im Säuglingsschrei von monozygoten Zwillingen eindrucksvoll belegen, dass die Entwicklung des Säuglingsschreis universalen Evolutionsprinzipien, wie dem *Prinzip der modularen Komposition von Komplexität* und dem *Prinzip der Wiederholung und Spezialisierung* folgt (Wermke, 1993, 1994; Wermke, 2002; Wermke & Friederici, 2004). Die Entwicklung komplexer biologischer Strukturen basiert nach dem Prinzip der modularen Komposition auf einfachen Elementen, welche aus angeborenen Vorstufen entstehen, die sich differenzieren und stabilisieren. Phänomenologisch spricht man von Reifungsvorgängen. Die noch unreifen Elemente werden dabei als Protomodule bezeichnet, ausgereifte Elemente als Module. Letztere werden als Bausteine im Gehirn gespeichert und stehen somit zur intentionalen Anwendung in fast jedem Kontext zur Verfügung. Dabei ist die Duplizierung eines Moduls die einfachste Form der Komplexitätssteigerung, während eine höhere Komplexität erst durch die Kombination verschiedener Elemente erreicht wird (Wermke, 2002; Wermke & Mende, 2011, 2016). In einigen Studien zur vorsprachlichen Entwicklung (siehe Kapitel 2) wurde die melodische Entwicklung auch in Komfortvokalisationen – ein Vokalisationsverhalten, welches direkt an die frühere Phase, in der Schreivokalisation überwiegen, anknüpft – untersucht. Dabei kamen die Autoren ähnlich wie bei der Melodie im Säuglingsweinen zu dem Ergebnis, dass die Komplexität der Melodie innerhalb von Komfortvokalisationen

in den ersten sechs Monaten kontinuierlich zunimmt. Vokalisationen mit einfachen Melodien überwiegen hier jedoch (Clement, 2004; Hsu et al., 2000; Kent & Murray, 1982; Legerstee, 1991; Oller, 1980; Stark, 1980). Die Tatsache, dass einfache melodische Strukturen in der hier untersuchten Vokalisationsphase überwiegen, gibt Anlass zu der Annahme, dass das universale Evolutionsprinzip der Wiederholung hinsichtlich der Entwicklung melodischer Strukturen hier auf anderer Ebene auch greift. Obwohl der Säugling in der Phase des Säuglingsschreis bereits komplexe Melodiemodule erworben hat, scheint er in der darauffolgenden Phase (Komfort) zunächst wieder auf einfache (Proto-) Module zurückzugreifen.

Eine Besonderheit der hier untersuchten und an die Phase des Säuglingsschreis anknüpfenden Phase der Komfortvokalisationen besteht in einem Zusammenschalten laryngealer und supralaryngealer Mechanismen. So wird der Grundfrequenzverlauf von Komfortvokalisationen – anders als in vorangegangenen Schrei-Vokalisationen – zunehmend durch artikulatorische Elemente – produziert auf suprasegmentaler Ebene – unterbrochen (siehe Kapitel 2). Neben melodischen und zeitlichen Eigenschaften entwickelt sich im ersten halben Lebensjahr – bedingt durch anatomische und neurophysiologische Reifungsvorgänge auf supralaryngealer Ebene (siehe Kapitel 1) - also auch die Produktion von artikulatorischen Elementen in den Vokalisationen der Säuglinge (Koopmans-van Beinum & van der Stelt, 1986; Oller, 1980; Stark, 1980). Dabei nimmt zum einen das Verhältnis von artikulierten zu nicht-artikulierten Vokalisationen zu, zum anderen verändern sich die Artikulationszonen im Vokaltrakt (Clement, 2004; Grenon et al., 2007; Irwin, 1946, 1947a, 1947b, 1951; Smith & Oller, 1981; Stoel-Gammon, 1988; Stoel-Gammon & Otomo, 1986).

Ausgehend von den Ergebnissen vorangegangener Arbeiten stellt sich in vorliegender Arbeit die übergeordnete Forschungsfrage, ob es im Untersuchungszeitraum zu einer Zunahme komplexer Strukturen auf laryngealer und supralaryngealer Ebene in Komfortvokalisationen kommt. Konkret sollen folgende spezifische **Forschungsfragen** beantwortet werden:

- Nimmt die *Komplexität der Melodiestructur* der Komfortvokalisationen über den Untersuchungszeitraum *kontinuierlich zu*?

- Nimmt die *Häufigkeit der Produktion artikulatorischer Elemente in Komfortvokalisationen kontinuierlich zu* und erfolgt eine *Verschiebung von der hinteren zur vorderen und mittleren Artikulationszone*?
- *Melodie als „Baugerüst“* für erste artikulatorische Bewegungen: Verläuft die Entwicklung auf laryngealer und supralaryngealer Ebene synchron?
- *Veränderungen der stimmphysiologischen Charakteristika* aufgrund von anatomischen und neurophysiologischen Reifungsvorgängen: Kommt es zu einem *Anstieg der mittleren Vokalisationslänge* sowie einem *Anstieg der Grundfrequenzeigenschaften (mittlere F_0 , F_0 -Range)* in den Komfort-Vokalisationen?

Studiendesign und Datenerhebung

Ausgehend von der Zielstellung der vorliegenden Arbeit wurde ein längsschnittliches Studiendesign gewählt, welches in 5 monatliche Erhebungszeiträume unterteilt wurde. Die Produktion erster Komfortvokalisationen fällt – in Folge rasanter neurophysiologischer und anatomischer Entwicklungsschritte – ca. in den Zeitraum zwischen dem 3. und 7. Lebensmonat (Expansionsphase). Diese Phase wurde daher für den Untersuchungszeitraum vorliegender Arbeit gewählt. Um einen hinreichend repräsentativen Lautkorpus untersuchen zu können wurde in Anlehnung an Erfahrungswerte am ZVES eine Mindestäußerungsmenge von 45 Komfortvokalisationen pro Kind und Monat gewählt. Um diese zu erhalten, wurde ein engmaschiger wöchentlicher Aufnahmeturnus angestrebt. Die Aufnahmen wurden von der Autorin im häuslichen Umfeld der Kinder in Anwesenheit einer vertrauten Bezugsperson angefertigt: Erfahrungen aus einem zeitgleich laufenden Datenerhebungsprojekt am ZVES haben gezeigt, dass es äußerst schwierig ist Kinder dieses Alters in ungewohnter (klinischer) Umgebung zu spontanen Komfort-Vokalisationen zu motivieren (Wermke K., Linz C., Hasenberg A., Kunz F., Meyer-Marcotty P., Schweitzer T., 2017). Allein die Anwesenheit einer fremden Person brachte die Kinder oft regelrecht zum „Verstummen“. Gründe dafür sind u.a. darin zu suchen, dass die Kinder mit ihren Äußerungen sozialen Interaktionspartnern emotionale Antworten/Reaktionen entlocken wollen, also in erster Linie ihren vertrauten Bezugspersonen (Oller et al., 2013). „They are produced when infants are alert and relaxed or playful“ (Gratier et al., 2015, p. 237). Die Angaben über

die Bereitschaft von Säuglingen im ersten halben Lebensjahr in einem bestimmten Zeitraum eine bestimmte Menge an Äußerungen zu vokalisieren variiert in unterschiedlichen Forschungsarbeiten sehr stark und kann zwischen 1 und 20 Vokalisationen pro Minute liegen. Beispielsweise konnte in manchen 20-minütigen Aufnahmesamples nicht eine einzige Äußerung aufgenommen werden, während zu einem anderen Aufnahmezeitpunkt bis zu 25 brauchbare Vokalisationen in der gleichen Zeit erfasst werden konnten (Michelsson et al., 2002). Dies bestätigte sich auch aus Erfahrungen am ZVES und musste bei der Planung der Datenerhebung berücksichtigt werden.

Um den Aufwand für die teilnehmenden Familien und die Autorin, welche allein für die Datenerhebung verantwortlich war, in einem angemessenen Rahmen zu halten, wurde daher entschieden die Aufnahmen teilweise von den Eltern selbst anfertigen zu lassen. Dadurch konnte sichergestellt werden, die Aufnahmen im Fall einer vorübergehenden geringen Vokalisationsbereitschaft zu unterbrechen und zu einer beliebigen Tageszeit, z.B. direkt nach dem Aufwachen, fortzusetzen. Dadurch sollte die Drop-out Rate möglichst geringgehalten werden. So entfielen längere Anfahrtswege an das ZVES (bis zu 15 km und weiter) für die Familien, wodurch der zeitliche und finanzielle Aufwand möglichst gering und die Bereitschaft zur Teilnahme über mehrere Monate möglichst hochgehalten werden konnte.

Um dennoch qualitativ hochwertige Aufnahmen realisieren zu können, musste vor Beginn der Datenerhebung eine Auswahl möglicher Aufnahmeorte im häuslichen Umfeld der Familien zusammen mit der Autorin erfolgen. So konnten mögliche Störquellen (z.B. elektrische Geräte) oder ungünstige akustische Raumbedingungen durch Probeaufnahmen, welche im Labor auf ihre Qualität überprüft wurden, identifiziert und ausgeschlossen werden. Des Weiteren mussten eine genaue Einweisung zur Bedienung der Geräte (z.B. Einsetzen des Speichermediums, Aufnahmeeinstellungen, Laden des Akkus, Mikrofoneinstellungen) sowie zum Ausfüllen geeigneter Aufnahmeprotokolle (Datum, Uhrzeit, Aufnahmeort, Besonderheiten der Aufnahmesituation) seitens der Autorin geplant und durchgeführt werden.

Um den Zeitraum, über den sich die Datenerhebung im Projekt erstrecken sollte, so kurz wie möglich zu halten, musste für die Planung der Datenerhebung zusätzlich die gleichzeitige Verfügbarkeit einer möglichst hohen Anzahl an digitalen Aufnahme-Recordern des ZVES über einen Zeitraum von mehreren Monaten sichergestellt werden. Obwohl im gleichen Zeitraum ein weiteres Datenerhebungsprojekt am ZVES durchgeführt wurde, welches ebenfalls den Einsatz mehrerer Aufnahmegeräte

erforderte, konnte im Rahmen vorliegender Arbeit ein Probandenkollektiv von 11 Kindern untersucht werden.

Probandenauswahl

Das kindliche Vokalisationsrepertoire in der Phase zwischen dem 3. und 7. Lebensmonat beschreibt sich, wie bereits aufgezeigt (siehe Kapitel 2), durch eine große inter- und intraindividuelle Variabilität. Letztere ist kennzeichnend für diese Phase und sollte daher mit Hilfe zeitlich dichter Aufnahmen erfasst werden. Die Ursachen der interindividuellen Unterschiede im Vokalisationsrepertoire, z.B. die Dauer der Phase oder das Auftreten/ die Qualität/ die Komplexität einzelner Vokalisationstypen innerhalb dieser Phase sind bisher nicht ausreichend erforscht. Umso wichtiger erschien es bei der Probandenauswahl einige Einflussfaktoren, die die frühe Sprachproduktion beeinflussen könnten, zu homogenisieren. Dazu zählen das Risiko für spätere Sprachentwicklungsstörungen, der Hörstatus, der muttersprachliche Input sowie der sozioökonomische Status der Eltern.

Rescola & Ratner (1996) untersuchten jeweils 30 zweijährige Kinder mit und ohne spezifischer Sprachentwicklungsstörung (SLI) und fanden heraus, dass die Kinder mit SLI deutlich seltener vokalisiert als die unauffälligen Kinder (Cierpka, 2014). Stoel-Gammon (1989) untersuchte in einer retrospektiven Studie bei zwei „Late Talkern“¹⁸ das Babbel-Inventar im Alter zwischen 9 und 24 Monaten und konnte ein limitiertes phonetisches Inventar sowie die Verwendung vereinfachter Silbenstrukturen feststellen (Stoel-Gammon, 1989). Da ähnliche Effekte in frühen Komfortvokalisationen im Alter zwischen 3 und 7 Monaten nicht auszuschließen sind, mussten diese Befunde u.a. bei der Probandenauswahl berücksichtigt werden. Mittels eines am ZVES erstellten Fragebogens wurden vor der Teilnahme an der vorliegenden Studie hereditäre Risikofaktoren, wie z.B. eine positive Familienanamnese bezüglich spezifischer Sprachentwicklungsstörungen bzw. angeborener Hörstörungen abgefragt. Kinder mit positiver Familienanamnese wurden von der Studie ausgeschlossen. Retrospektiv erfolgte im Alter von 12 und 18 Monaten die Ausgabe standardisierter Elternfragebögen zur Identifizierung von Risikokindern für eine spätere SSES (Grimm & Doil, 2000).

Oller und Kollegen zeigten, dass ein geringer bzw. sehr geringer sozioökonomischer Status der Eltern ebenfalls Einfluss auf das Vokalisationsverhalten haben können (Oller et al., 1995). Die Autorin entschied daher, Kinder aus Familien mit geringem oder sehr

¹⁸ Kinder die spät in die produktive Sprache einsteigen und mit 24 Monaten weniger als 50 Wörter sprechen und noch keine 2-Wort-Sätze bilden (Dannenbauer, 2009).

geringem sozioökonomischem Status (Abfrage von Schulabschluss, Beruf, Einkommen der Eltern) von der Studie auszuschließen.

Um potenzielle Effekte einer zweiten, nicht-deutschen Muttersprache ausschließen zu können, wurden auch bilingual aufwachsende Kinder von der Studie ausgeschlossen. Die genauen Angaben zur Probandenauswahl können ausführlich in Kapitel 5.1 nachgelesen werden.

Lautauswahl

Wie bereits beschrieben, ist das Ziel der vorliegenden Arbeit die Charakterisierung von Komfortvokalisationen monolingual deutsch aufwachsendender Kinder zwischen dem 3. und 7. Lebensmonat, insbesondere die Beschreibung deren melodischer und artikulatorischer Eigenschaften. Damit fokussiert sich die Untersuchung auf spontane Lautäußerungen, die aufgrund auditiver und spektraler Eigenschaften weder der vorangehenden Phase (Schrei) noch der darauffolgenden Phase (Kanonisches Babbeln) zugeordnet werden können. Da jedoch – wenn auch in begrenzter Anzahl – andere Vokalisationstypen in dieser Phase auftreten, musste eine Vor-Analyse des Vokalisationsrepertoires der Aufnahmen durch die Autorin geplant werden. Wie die Identifikation von Komfortvokalisationen (Voranalyse) für die spezifische Datenanalyse erfolgte ist in Kapitel 5.5.1 beschrieben.

Wie in Kapitel 2 ausführlich beschrieben, wird ab dem 2. Lebensmonat eine ganze Bandbreite an Vokalisationstypen produziert: Zunächst sind neben rein vokalischen Lauten, sog. „*quasi-resonant nuclei*“¹⁹, auch *Cooing-Laute*²⁰ zu erwarten, welche scheinbar freiwillig und oft als Antwort auf soziale Interaktionen einer Bezugsperson produziert werden²¹. In der Phase des *Vocal Play* ab dem 4. Lebensmonat erlangt das Kind zunehmend mehr Kontrolle über laryngeale und artikulatorische Mechanismen. Es werden zum einen prosodische Eigenschaften wie die Tonhöhe und deren Variation (resultiert in sog. „squeals“ und „growls“) sowie die Lautstärke (resultiert in sog. „yells“ und „whisper“) manipuliert. Zum anderen werden artikulatorische Eigenschaften manipuliert. Dies resultiert v.a. in Reibelauten („friction noises“), nasalen Murren („nasal murmurs“) und bilabialem und uvularem „Trillern“ („raspberries“ und „snorts“). Des Weiteren werden nun häufiger erwachsenähnliche Vokale („adult-like“ vowels)

¹⁹ Beschrieben durch Oller: Vokalisationen mit normaler Phonation jedoch begrenzter Resonanz, produziert mit (fast) geschlossenem Mund, nasal auditiver Eindruck (Vihman, 1996).

²⁰ Laute mit konsonantenähnlicher (frikativer) Überlagerung, typischerweise velar; Einzel- oder Serienlaute getrennt durch glottal stops (Vihman, 1996).

²¹ Die im Folgenden aus dem Englischen übernommenen Begriffe werden ausführlich in Kapitel 2 (siehe Fußnoten 8-12) beschrieben.

sowie marginales Babblen (hier werden konsonantische und vokalische Elemente verbunden) produziert (Vihman, 1996). In den wenigen Studien, die vorsprachliche Komfortvokalisationen im Alter zwischen 3 und 7 Monaten beschreiben, wurden oftmals nur bestimmte Vokalisationstypen (z.B. „sprach-ähnliche Laute“, marginale Babbell-Laute) untersucht bzw. wurden bestimmte Vokalisationstypen ausgeschlossen, z.B. „squeals“, „growls“, „yells“, „raspberries“ (Rothgänger, 2003; Scheiner et al., 2002).

Gemäß der Zielstellung der vorliegenden Arbeit, wurde dieser Ansatz jedoch verworfen. Die Autorin der vorliegenden Arbeit hat sich bewusst gegen einen vorab Ausschluss bestimmter Vokalisationstypen im Repertoire der Komfortvokalisationen entschieden. D.h. es wurden alle Nicht-Schrei-Vokalisationen in die Analyse eingeschlossen. Ein Ausschluss bestimmter Vokalisationstypen beinhaltet das Risiko charakteristische Veränderungen und Besonderheiten der hier untersuchten Phase zu ignorieren bzw. nicht zu bemerken. Für eine objektive Charakterisierung der Expansionsphase im Alter von 3 bis 7 Monaten sollten alle auftretenden Vokalisationstypen in die Analyse eingeschlossen werden. Auch diesbezüglich stellt die vorliegende Arbeit Neuland dar. Bisher liegen keine vergleichbaren Untersuchungen in der Fachliteratur vor.

Analyseparameter und Datenauswertung

Wie bereits erwähnt, knüpft die „Komfort-Phase“ direkt an die Phase des Schreis an. Es ist davon auszugehen, dass sich die neurophysiologischen und anatomischen Reifungsvorgänge in dieser Phase u.a. in Veränderungen der stimmphysiologischen Parameter der Komfortvokalisationen widerspiegeln. Daher sollte die mittlere Grundfrequenz, der Range (Hub) der mittleren Grundfrequenz innerhalb einer Vokalisation sowie die Vokalisationslänge im Längsschnitt erfasst werden. Die Analyseschritte- und Verfahren zur Erfassung der stimmphysiologischen Parameter werden ausführlich in Kapitel 4 beschrieben.

Um feststellen zu können, ob sich die Melodie in Komfortvokalisationen, ähnlich wie in Schrei-Vokalisationen, nach einem systematischen Muster entwickelt, sollte in einem weiteren Analyseschritt auch die Melodiestructur pro Vokalisation erfasst werden. Dabei wurden die einzelnen Vokalisationen einfachen oder komplexen Melodiestructurkategorien zugeordnet. Auf Basis dieser Strukturkategorien konnte ein Melodiekomplexitätsindex (MCI) pro Aufnahmezeitpunkt (pro Kind) ermittelt werden. Dieser MCI gibt im Längsschnitt Aufschluss über den Komplexitätsgrad der untersuchten Komfortvokalisationen. Die dem MCI zugrundeliegenden verwendeten

Strukturkategorien und das geeignete Analyseverfahren standen der Autorin im ZVES zur Verfügung und werden in Kapitel 4 detailliert beschrieben.

Neben der larynx-fokussierten Analyse sollte auch die Erfassung der Parameter auf supralaryngealer Ebene – also die Analyse der artikulatorischen Aktivität erfolgen. Eine Möglichkeit zur Analyse der artikulatorischen Fähigkeiten besteht in der Beschreibung produzierter konsonantischer Elemente sowie der Untersuchung deren Auftretshäufigkeit im Untersuchungszeitraum (Robb et al., 2020). Dabei hat sich die Autorin bewusst gegen eine Orientierung an der Erwachsenensprache – wie sie partiell auch bei dem infraphonologischen Ansatz nach Oller vorkommt – entschieden (Oller, 1986). In der Vokalisationsphase im Alter zwischen 3 und 7 Monaten werden neben den laryngealen v.a. auch die supra-laryngealen Resonanzräume explorativ erkundet. Dies spiegelt sich u.a. in der Verwendung eines unbegrenzten Repertoires an artikulatorischen Elementen in Komfortvokalisationen wider. Dieses Repertoire gilt als universal, d.h. für die Verwendung von segmentalen Eigenschaften liegen (noch) keine linguistischen Begrenzungen durch die Zielsprache vor (Vihman, 1996). Um diesen explorativen Charakter der untersuchten Vokalisationsphase besser erfassen zu können sollten die artikulatorischen Eigenschaften nicht mittels einer an der Zielsprache orientierten phonetischen Transkription, sondern lediglich durch die Kategorisierung von Artikulationszonen und Artikulationsarten (konsonantisch vs. vokalisch) beschrieben werden.

Dabei musste zunächst zwischen konsonantischen Artikulationen und Artikulationen aufgrund von Resonanzveränderungen (Vokaländerung innerhalb einer Vokalisation) unterschieden werden. Bei Letzteren handelt es sich nicht um artikulatorische Elemente im engeren Sinne, die aufgrund von Engebildungen oder Verschlüssen in der Mundhöhle (konsonantisch) zustande kommen, sondern um vokalische Artikulationen, die aufgrund von Veränderungen der Vokaltraktgeometrie (Resonanzveränderungen) während der Phonation zustande kommen (Wirth, 2000). Auch für vokalische Veränderungen ist eine Koordination der Sprechwerkzeuge auf supralaryngealer Ebene verantwortlich. Aus diesem Grund wurden diese bei der Analyse miterfasst (siehe Kapitel 4).

In der explorativen Datenanalyse sollte zunächst die relative Anzahl artikulierter Laute im Untersuchungszeitraum dargestellt werden. Da dies noch keinen Rückschluss über die Komplexität der artikulatorischen Leistungen zulässt, wurde im nächsten Schritt die Anzahl der artikulatorischen Elemente pro Vokalisation ermittelt. Anschließend wurde die Verteilung der artikulatorischen Elemente hinsichtlich verschiedener Artikulationszonen im Verlauf untersucht, nur so lässt sich eine potenzielle Verschiebung

erkennen (siehe Kapitel 4). Da davon auszugehen ist, dass diese Entwicklung von individuellen Unterschieden geprägt ist, sollte diese Analyse nicht nur auf Gruppenebene, sondern auch auf Kindebene durchgeführt werden.

Neben der Analyse auf Vokalisationsebene erfolgte im letzten Schritt die Analyse der artikulierten Laute auf Melodiestrukturbene. So konnte untersucht werden ob bzw. wie sich die Artikulationsentwicklung in Abhängigkeit von der Melodieentwicklung vollzieht. Hierzu zählte im letzten Schritt die Verteilung der artikulierten Laute über die einzelnen Melodiestrukturen in Abhängigkeit vom Alter.

Die Datenauswertung sollte im ersten Schritt auf Vokalisationsebene und Kindebene mittels deskriptiver Statistik erfolgen. Für die statistischen Analysen auf Kindebene sollten im Anschluss nichtparametrische Testverfahren Anwendung finden (siehe Kapitel 4).

4. Material und Methoden

In der vorliegenden explorativen Längsschnittuntersuchung wurden die phonatorischen und artikulatorischen Eigenschaften spontaner Komfortvokalisationen bei Säuglingen im Alter zwischen 3 und 7 Monaten analysiert. Die Durchführung erfolgte im Rahmen einer Zusammenarbeit zwischen dem Lehrstuhl für Sprachheilpädagogik der Universität Würzburg (Prof. Dr. Detlef Hansen) und dem Zentrum für vorsprachliche Entwicklung und Entwicklungsstörungen (ZVES; Leitung: Prof. Dr. Kathleen Wermke) sowie dem Comprehensive Hearing Center (CHC) der HNO-Klinik des Universitätsklinikum Würzburg (Prof. Dr. med. Wafaa Shehata-Dieler) auf dem Gebiet der Erforschung vorsprachlicher Entwicklungsverläufe. Die Lautaufnahmen erfolgten im Rahmen einer am Universitätsklinikum (Medizinische Fakultät) durchgeführten Studie, welche durch das Ethik-Votum (145/12) im Juli 2012 beraten wurde. Die Aufnahmen erfolgten in der häuslichen Umgebung der Kinder im näheren Umkreis von Würzburg (siehe 4.2). Vor Studienbeginn erhielten die Eltern eine ausführliche Information zum Projekt und unterzeichneten eine schriftliche Einwilligungserklärung zur Teilnahme und Nutzung der Spontansprachdaten für wissenschaftliche Zwecke. Die Aufbereitung und Analyse der Aufnahmen erfolgte im Signalanalyselabor am ZVES, da hier die erforderlichen technischen Voraussetzungen gegeben waren und eine spezielle Analysesoftware (siehe 4.3, 4.4) zur Verfügung stand.

4.1 Probanden

Im Rahmen der Längsschnittuntersuchung wurden spontansprachliche Vokalisationen von 11 gesunden peri- und postnatal unauffälligen Säuglingen (5 männlich/ 6 weiblich) im Alter zwischen 64 und 203 Tagen erhoben. Sie wurden alle termingerecht geboren und wachsen ausschließlich deutschsprachig auf. Der Kontakt erfolgte zum einen über Aushänge in Hebammenpraxen und Kindergärten in Würzburg über den Projektflyer und zum anderen durch persönliche Kontakte im näheren Bekanntenkreis der Autorin. Die Familien wurden in regelmäßigen Abständen besucht. So konnte die Gesamtentwicklung der Kinder verfolgt werden. Der sozioökonomische Status der Herkunftsfamilie ist bei allen Probanden als mittel bis hoch zu bezeichnen. Es lagen

keine hereditären Risikofaktoren wie z.B. eine positive Familienanamnese bezüglich spezifischer Sprachentwicklungsstörungen bzw. angeborener Hörstörungen vor.²²

Um etwaige Kinder mit einem Risiko für spätere Sprachentwicklungsstörungen retrospektiv von der Analyse ausschließen zu können, wurde der ELFRA-1 (Grimm & Doil, 2000) mit 12 und 18 Monaten herausgegeben. Während der ELFRA-1 mit 12 Monaten bei zwei Kindern (VF, VH) im Bereich Sprachproduktion auffällig war, zeigte die Durchführung im Alter von 18 Monaten nur noch Auffälligkeiten bei einem Probanden (VF). Dieser Proband zeigte auch bei der Durchführung des SETK-2 einen unterdurchschnittlichen Wert im Bereich Produktion (siehe Tabelle 31, Anhang) und muss daher unter Umständen gesondert diskutiert werden. Es kann davon ausgegangen werden, dass unter den restlichen Probanden der vorliegenden Studie keine Risikokinder für spätere Sprachentwicklungsstörungen sind.

Temporäre Beeinträchtigungen des Hörvermögens, wie z.B. durch Mittelohrentzündungen oder Paukenergüsse, sind im Untersuchungszeitraum nicht wesentlich aufgetreten. Die Ergebnisse der Vorsorgeuntersuchungen (U1 und U2) waren bei allen Kindern unauffällig.

4.2 Datenerhebung - Lautaufnahmen

Die Lautaufnahmen erfolgten in regelmäßigen Abständen von 10-14 Tagen zwischen dem 64. und 203. Tag (siehe **Tabelle 1**) mit einem akkubetriebenen digitalen Aufnahmerecorder der Marke Tascam DR-100 (Samplingfrequenz 48 kHz/ Dynamikbereich 16 Bit) sowie einem omnidirektionalen Kondensatormikrofon der Marke Earthworks TC-20 unter manueller Aussteuerung. Als Speichermedium diente eine herausnehmbare SD-Speicherkarte, wodurch die Lautaufnahmen direkt im *.wav-Format gespeichert wurden.

Während der Aufnahmen betrug der Abstand zwischen Mikrofon und Schallquelle (Mund des Kindes) etwa 15 cm. Sämtliche Aufnahmen erfolgten in der häuslichen Umgebung der Kinder. Sie wurden von der Autorin selbst oder von den Eltern angefertigt. Dabei wurde der jeweilige Kontext der Aufnahme sowie Datum, Uhrzeit und relevante Angaben zum Gesundheitszustand des Kindes auf der Aufnahme festgehalten. Zusätzlich wurden Aufnahmeprotokolle angelegt, auf denen ergänzend Angaben zum Hörstatus, der Nahrungsaufnahme und dem allgemeinen Entwicklungsstand (z.B. motorische

²² Erfassung mittels eines am ZVES erstellten Fragebogens.

Entwicklung) des Kindes vermerkt wurden. Diese Angaben dienten nur der Dokumentation eines unauffälligen Entwicklungsverlaufs.

Die Aufnahmedauer hing immer von der Kommunikationsbereitschaft bzw. Lautierfreudigkeit der Kinder ab. Ziel war es spontane Lautäußerungen in möglichst natürlichen Interaktionssituationen mit einem oder beiden Elternteilen zu erheben. Es wurde bei jeder Aufnahme abgewartet, bis das Kind von sich aus begann spontane Komfortvokalisationen zu äußern. Die Aufnahmedauer variierte in der Regel zwischen 30 und 60 Minuten. Ziel war es zu jedem Aufnahmezeitpunkt mindestens 30 nicht vegetative spontane Lautäußerungen in Folge zu erhalten. Die Aufnahmesequenz wurde im Abstand von 14 Tagen angesetzt. Konnte die Mindestzahl von 30 nicht vegetativen spontanen Lautäußerungen zu einem Aufnahmezeitpunkt nicht erreicht werden, wurden innerhalb des 14-tägigen Zeitraums weitere Aufnahmen angefertigt, sodass in den meisten Fällen ein wöchentlicher Aufnahmeturnus entstand. Es wurde eine Mindestzahl von 45 Lauten pro Monat pro Säugling angestrebt. Dadurch variiert die Anzahl der Aufnahmezeitpunkte zwischen den einzelnen Kindern teilweise stark. Ein weiterer Grund für die unterschiedliche Anzahl der Aufnahmezeitpunkte pro Kind besteht im Auswahlkriterium der zu analysierenden Vokalisationstypen. Aufnahmezeitpunkte, bei denen die Anzahl kanonischer Vokalisationen überwiegte, wurden von der Analyse ausgeschlossen. Grund dafür ist die Annahme, dass in diesem Fall bereits eine fortgeschrittene Phase des Vokalisationserwerbs erreicht ist (siehe Lautauswahl). Da der Zeitpunkt des Eintritts in diese Phase zwischen den einzelnen Kindern variierte, variierte auch die Anzahl der Aufnahmezeitpunkte. Insgesamt wurden 127 aufgezeichnete Kommunikationssituationen für die anschließende Analyse ausgewählt (siehe Tabelle 1).

Tabelle 1: Probandentabelle mit Aufnahmezeitpunkten pro Kind

Kind	Alter in Tagen										
AC	84	91	99	118	152						
AD	64	79	107	111	115	119	124	128			
AK	82	89	91	108	120	128	140	172	175		
VB	65	87	102	104	120	134	142	154	162	171	180
	191										
VD	80	94	105	121	142	149	150	154	157	161	173
	183										
VE	82	96	105	126	145	159	167	179			
VF	79	82	83	99	100	107	113	118	128	143	151
	174										
VG	74	81	91	97	107	120	127	138	148	157	169
	178	185	195								
VH	87	95	104	110	111	117	121	122	124	135	141
	146	149	151	156	164	175	176	181	187	193	
VI	73	80	89	97	102	108	118	130	141	158	160
	180	184									
VL	80	90	96	102	125	126	137	139	141	173	182
	195	203									

4.3 Datenaufbereitung

Für die eigentliche Analyse mussten die digitalen Lautaufnahmen zunächst in einen mit dem professionellen Sprachanalyseprogramm *Computerized Speech Lab* (CSL™, Modell 4500, KayPENTAX, USA) ausgestatteten PC am ZVES eingelesen werden. Mit Hilfe der zugehörigen Software wurden die 30- bis 60-minütigen Aufnahmen in einzelne, besser handhabbare Langserien (120s) geschnitten und mit einem pseudonymisierten Filenamen versehen gespeichert. Auf entsprechend vorliegenden Digitalisierungsprotokollen wurden Aufnahmesituation, Stimmung sowie sonstige Besonderheiten (z.B. Hintergrundgeräusche, Unterbrechung der Aufnahme durch Stillen/Füttern) schriftlich dokumentiert.

Anschließend wurden die selektierten Einzelvokalisationen innerhalb einzelner Expirationsphasen in sog. Einzelsignale geschnitten und gegebenenfalls verstärkt. Zusätzlich wurden diese je nach Äußerungsart unterschiedlich codiert (c=Schrei, b=Babbeln, m=mitigiert, n=non-cry) und mit dem jeweiligen Grad der Verstärkung im Digitalisierungsprotokoll vermerkt. Dies vereinfachte später die Auswahl der zu analysierenden Vokalisationen, da lediglich Komfortvokalisationen (Babbeln) in die Analyse einbezogen wurden (siehe 4.5.1.).

In einem ersten Analyseschritt mittels automatischer CSL-Routine im ZVES-Analysesystem wurde von jeder Einzelvokalisation ein Frequenzspektrogramm (Schmalband) erstellt (FFT, Hanning-Fenster, 1024 Abtastpunkte, Bandbreite 45 Hz), siehe **Abbildung 1**.

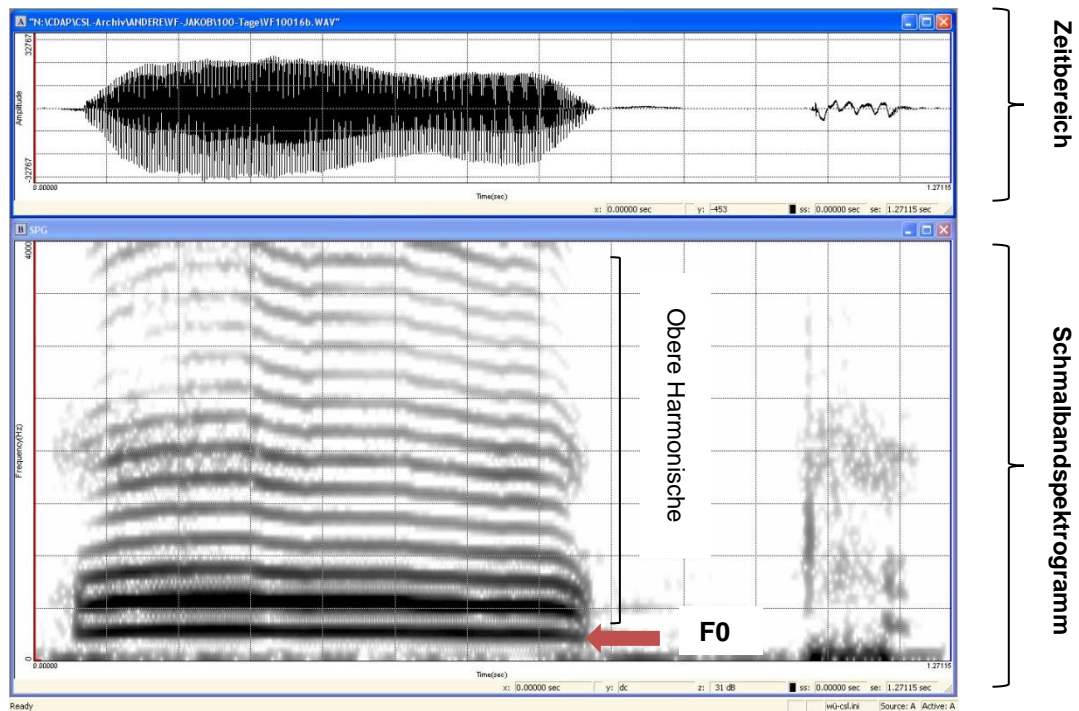


Abbildung 1: Schmalbandspektrogramm (unteres Fenster) mit korrespondierender Zeitbereichsdarstellung (oberes Fenster) einer Einzelvokalisation

Im oberen Fenster ist die Einzelvokalisation in der Zeitbereichsdarstellung abgebildet. Im unteren Fenster ist das instationäre Frequenzspektrum zu sehen. Auf der x-Achse des Spektrogramms ist der Zeitverlauf in Sekunden (s) und auf der y-Achse der Verlauf der Frequenz linear bis 4000 Hertz (Hz) dargestellt. Die Grundfrequenz F_0 wird durch den untersten Balken repräsentiert (siehe roter Pfeil). Ihre Zeitfunktion wird im Folgenden als Melodie bezeichnet. Sie spiegelt die zeitveränderlichen Schwingungen der Stimmlippen im Larynx wider. Alle darüber liegenden „Frequenzbänder“ sind die ganzzahligen Vielfachen der Grundfrequenz (Harmonische). Die Grauskalierung des Spektrogramms gibt Aufschluss über die Intensität des Signals; je dunkler eine Harmonische ist, desto mehr Energie enthält sie. Mit Hilfe des Spektrogramms können zum einen alle relevanten spektralen Eigenschaften, wie Subharmonische, phonatorisches Rauschen, Frequenzsprünge (shift) und zum anderen Artikulationsbewegungen erkannt und beurteilt werden.

4.4 Berechnung der Grundfrequenz- und Intensitätsverläufe

Die Berechnung der Melodie- und Intensitätsverläufe erfolgte automatisch mit Hilfe eines am ZVES erstellten Skripts (Skript-Autor: Annette Prochnow) der Open-Source Software *Praat* (Boersma & Weenink, 2009). Die *Praat-Software* ermöglicht die Anwendung einer Vielzahl an modernen Signalanalysemethoden und hat sich im Bereich der Sprach- und Kommunikationswissenschaften aufgrund seiner benutzerfreundlichen Oberfläche als Analyse-Werkzeug gegenüber anderen Systemen durchgesetzt. Mittels des am ZVES verwendeten Skripts war es möglich, eine große Anzahl an Signalen auf einmal (in einem Schritt) einzulesen und Grundfrequenz- und Intensitätswerte automatisch berechnen zu lassen. Für die automatische Berechnung der F_0 -Verläufe musste ein relevanter Frequenzbereich vorgegeben werden. Im Unterschied zur *CSL-Software* (Computerized Speech Lab) können mit der *Praat-Software* auch Grundfrequenzwerte die größer als 1000 Hz sind, berechnet werden. Dies war im Untersuchungszeitraum vorliegender Arbeit von großer Bedeutung, da in der ersten Hälfte des ersten Lebensjahres viele hochfrequente Vokalisationen auftreten. Durch Phänomene wie Frequenzsprünge, Subharmonische und Hintergrundgeräusche musste jede berechnete Melodiekurve mit dem F_0 -Verlauf im Schmalbandspektrogramm verglichen werden und gegebenenfalls unter Verwendung eines individuell angepassten Frequenzbereichs interaktiv mit Hilfe der *Praat-Software* von der Autorin neu berechnet werden.

4.5 Spezifische Datenanalyse

Ziel der spezifischen Datenanalyse ist die Erfassung der phonatorischen Eigenschaften (Grundfrequenz, F_0 -Range, Vokalisationslänge), sowie die Beschreibung der Melodiestructur und deren Verbindung mit ersten Artikulationsbewegungen in Komfortvokalisationen zwischen dem 3. und 7. Lebensmonat. Dazu müssen die Einzelvokalisationen sowohl quantitativ (Vokalisationslänge, F_0 -Hub, mittlere F_0) als auch qualitativ (Lautkategorisierung, Melodiestructur, Artikulationsbewegungen pro Bogen/Vokalisation) analysiert werden. Grundlage hierfür ist zum einen das am ZVES entwickelte Programm CDAP© der Firma „pw-project“ und zum anderen ein von der Autorin entworfenes Kodierungssystem zur Erfassung verschiedener Artikulationsbewegungen pro Vokalisation/Melodiebogen, siehe **Abbildung 6**.

4.5.1 Identifizierung von Komfortvokalisationen

Insgesamt wurden aus den an 127 Aufnahmetagen aufgezeichneten Sequenzen 7375 Einzelvokalisationen manuell segmentiert und pseudonymisiert archiviert. Wie bereits beschrieben, sollten in die spezifische Datenanalyse ausschließlich Komfortvokalisationen (z.B. Quasivokale, Gurren, marginales Babbeln) einfließen. Vokalisationen, die aufgrund ihrer spektralen und artikulatorischen Eigenschaften nicht diesem Vokalisationstyp zugeordnet werden konnten, wurden aussortiert. Dazu zählten vegetative Vokalisationen, Schrei-Vokalisationen, und Vokalisationen die bereits dem kanonischen Stadium zugeordnet werden konnten (kanonisches Babbeln). Aus dem Vokalisations-Repertoire von insgesamt 7375 Vokalisationen konnten 5476 geeignete Komfortvokalisationen für die spezifische Datenanalyse identifiziert werden (siehe **Abbildung 2**).

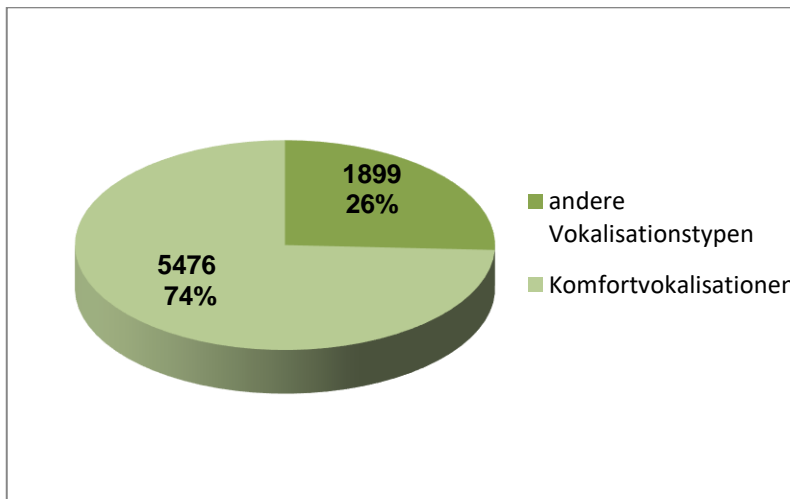


Abbildung 2: Zusammensetzung des Vokalisations-Repertoires

Die Datenanalyse erfolgte aus zwei Perspektiven und somit vokalisationsbasiert und kindbasiert. D.h. zum einen erfolgte die statistische Auswertung der Messgrößen auf Basis der Gesamtzahl aller Komfortvokalisationen und zum anderen auf Basis der Gesamtzahl der Komfortvokalisationen pro Kind. Die vokalisationsbasierte Analyse wurde zusätzlich pro Monat (3.-7. Monat) durchgeführt. **Tabelle 2** gibt Aufschluss über die Anzahl der analysierten Vokalisationen pro Kind/pro Monat.

Tabelle 2: Anzahl der analysierten Vokalisationen pro Kind/pro Monat

Anzahl der Vokalisationen		Alter in Monaten					Gesamt
		3	4	5	6	7	
Kind	AC	65	232	0	11	0	308
	AD	117	206	49	0	0	372
	AK	132	138	97	62	0	429
	VB	80	153	51	153	29	466
	VD	30	125	190	159	44	548
	VE	51	55	111	174	0	391
	VF	120	318	98	98	0	634
	VG	55	257	204	256	128	900
	VH	24	133	111	91	85	444
	VI	147	152	120	86	4	509
	VL	42	60	206	63	104	475
Gesamt		863	1829	1237	1153	394	5476

Wie bei der Anzahl der Aufnahmezeitpunkte pro Kind, zeigt sich auch bei der Anzahl der analysierten Vokalisationen pro Kind/pro Monat teilweise eine starke Variation. Die Unterschiede sind in erster Linie auf unterschiedliches Vokalisationsverhalten zurückzuführen. Während einige Kinder in den Aufnahmesituationen große Vokalisationsfreude zeigten, war es bei anderen Kindern teilweise schwer spontane Komfortvokalisationen während der Aufnahmezeit zu evozieren.

Wie bereits in Kapitel 4.2 beschrieben, unterschied sich auch der Zeitpunkt für den Eintritt in die kanonische Phase der einzelnen Kinder stark. Dadurch lassen sich v.a. die Unterschiede in der Vokalisationsanzahl im 7. Monat erklären. Zum Untersuchungszeitpunkt des 7. Monats gehen nur noch von insgesamt 6 Kindern Vokalisationen in die Analysen ein. Das Lautrepertoire der restlichen 7 Kinder bestand zu diesem Zeitpunkt bereits hauptsächlich aus kanonischem Babbeln, welches von den akustischen Analysen ausgeschlossen wurde. Daher ist die Gesamtzahl der Vokalisationen im 7. Monat wesentlich kleiner als zu den restlichen Untersuchungszeitpunkten zwischen dem 3. und 6. Monat. Eine der analysierten Vokalisationen wurde zu einem Großteil mit dorsaler Artikulation produziert. Dies hatte zur Folge, dass es bei der Berechnung der F_0 -Daten zu Fehlern kam, welche auch durch mehrfache Änderungen der Settings in Praat nicht korrigiert werden konnten. Somit stand diese Vokalisation nur zur Analyse der Artikulation und der Äußerungslänge zur Verfügung.

Weitere Gründe für die Variation der Anzahl der analysierten Vokalisationen pro Kind/pro Monat sind private Gründe der Familien, die keine Aufnahmen erlaubten, wie z.B. kleine Erkältungskrankheiten oder Urlaubsreisen.

4.5.2 Melodieanalyse

Die mit *Praat* berechneten Grundfrequenzverläufe konnten durch die in CDAP integrierte Oberfläche des Melodie-Intensitäts-Diagramms dargestellt und analysiert werden. Mithilfe des CDAP-Programms wurde die Melodie zunächst auf Korrektheit geprüft. Algorithmus-Artefakte in Form von spike-artigen Werten bzw. Artefakte durch Artikulationsbewegungen (z.B. bei Plosiven) innerhalb des Grundfrequenzverlaufs wurden gelöscht. Anschließend wurden die Melodien einer Gauß-Tiefpaß-Filterung (40 Hz) unterzogen. Des Weiteren wurden – falls vorhanden – Inspirationsgeräusche gelöscht. **Abbildung 3** zeigt die graphische Darstellung des Melodieverlaufs einer Vokalisation innerhalb der Benutzeroberfläche von CDAP.

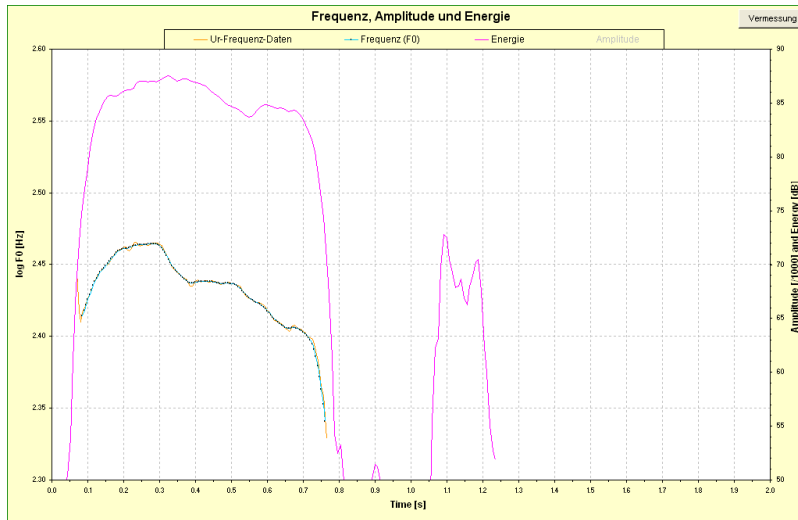


Abbildung 3: Beispiel für ein Melodie-Intensitätsdiagramm einer Vokalisation im CDAP-Programm. Die Dimensionsangaben erfolgen programmbedingt englisch.

Der blaue Graph zeigt den F_0 -Verlauf, also die Melodie innerhalb eines Lautes. Ein Anstieg dieser Linie spiegelt ein schnelleres Schwingen der Stimmlippen, also einen Anstieg der Melodie wider. Verlangsamt sich die Schwingung der Stimmlippen, fällt die Melodie ab. Auf der y-Achse kann die Einteilung der Frequenz in log Hertz (Hz) abgelesen werden. Der rote Graph stellt die Intensität des Signals dar, deren lautinterne Veränderung auf der rechten z-Achse in Dezibel (dB) abgelesen werden kann. Auf der x-Achse ist die Zeit [s] dargestellt.

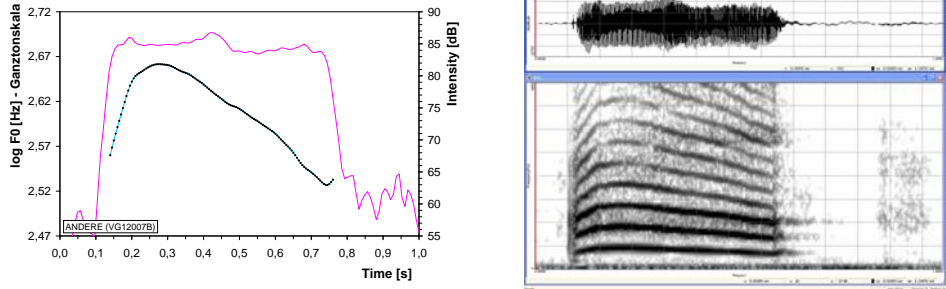
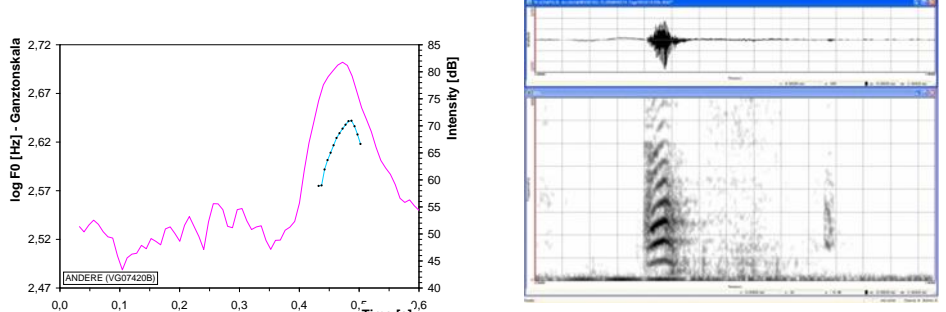
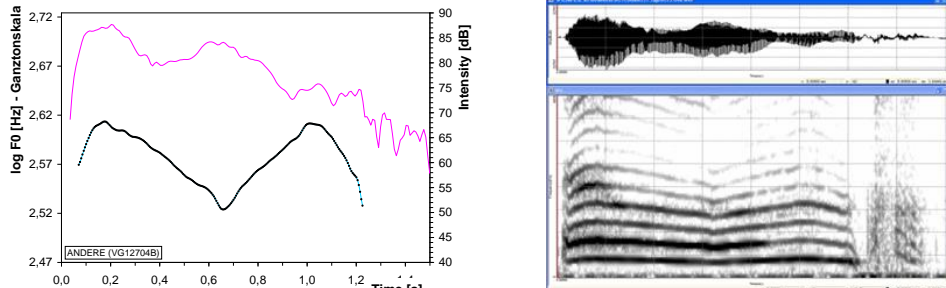
Angepasst an die menschliche Tonhöhenempfindung (Weber-Fechner-Gesetz), wurde für die Frequenzachse stets eine logarithmische Skala (log) gewählt, die in musikalische Halb- und Ganztonschritte unterteilt wurde. Um eine objektive Beurteilung der Melodiekontur zu garantieren wurden zur Analyse Standardeinstellungen der Skalen verwendet. Es wurde in den meisten Fällen eine einheitliche Zeitbasis von 2 Sekunden, eine Frequenzspannweite von 0,3 log Hz gewählt. Davon wurde nur abgewichen, wenn eine einzelne Vokalisation diese Werte überschritten hat.

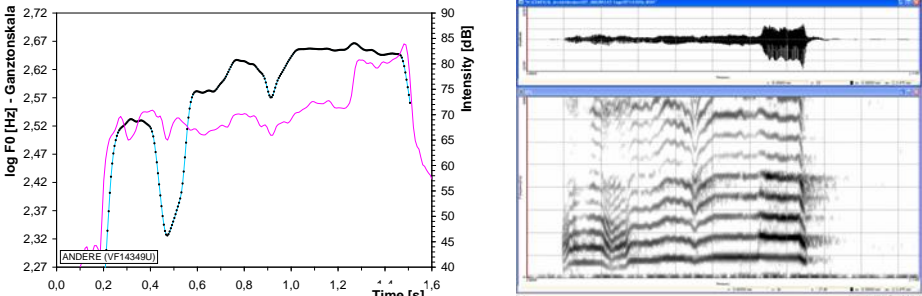
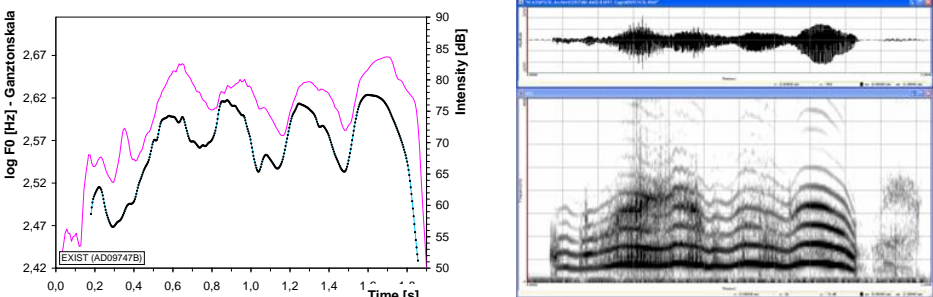
4.5.3 Melodiestrukturanalyse

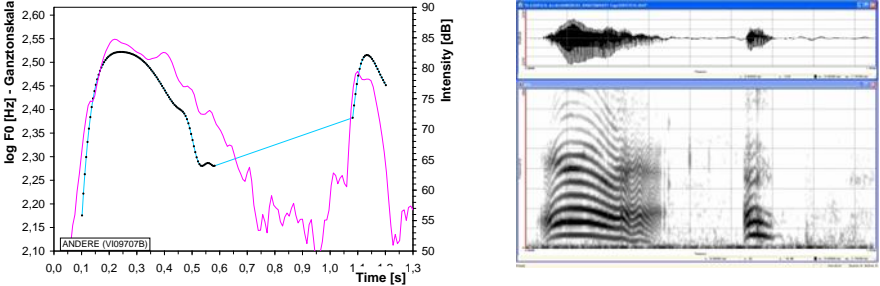
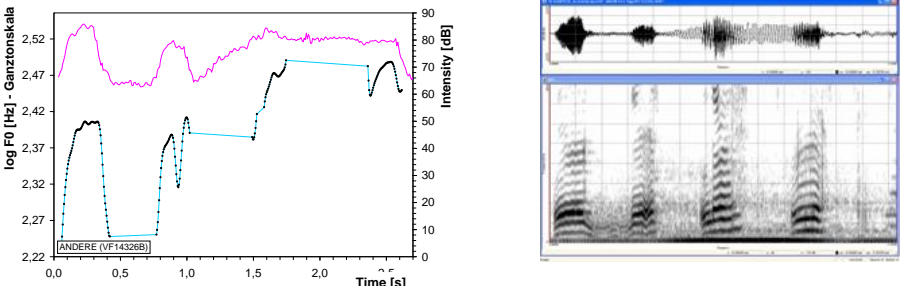
Ein Ziel der Analyse war es, die Melodiestruktur der Komfortvokalisationen im Alter zwischen dem 3. und 7. Monat systematisch im Längsschnitt zu erfassen. Grundlage für die Erfassung der Melodiestruktur bildeten dabei die im Rahmen der Deutschen Sprachentwicklungsstudie von Wermke und Kollegen (2007) entwickelten Strukturkategorien, welche in ein geeignetes Analysemodul in *CDAP* integriert wurden. Anhand der Schmalbandspektrogramme sowie der Melodie-Intensitätsdiagramme wurde unter auditiver Kontrolle die Melodiestruktur jeder Komfortvokalisation mithilfe eines speziellen Analysemoduls in *CDAP* analysiert und kategorisiert. Im Folgenden werden die Kriterien für die in dieser Arbeit verwendeten Melodiestrukturkategorien näher erläutert (siehe **Tabelle 3**).

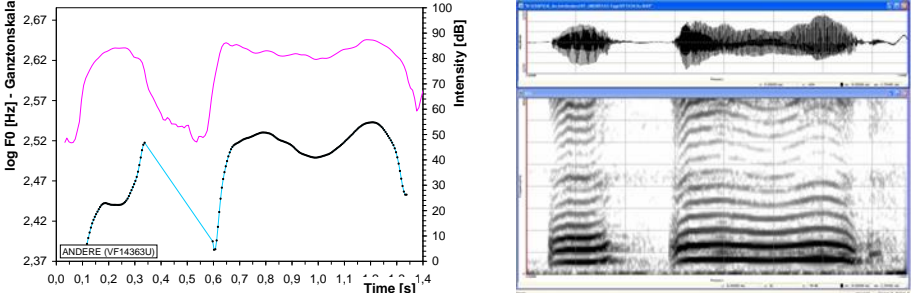
Für die finale Analyse der Melodiestrukturen im Entwicklungsverlauf erfolgte eine Zusammenfassung der einzelnen Strukturen in drei Kategorien. Kurzlaute und Melodiestrukturen mit einem Bogen wurden zu der Kategorie **1B** zusammengefasst. Unter der Kategorie **MB** wurden alle nicht segmentierten komplexen, also doppel- und mehrböjige Strukturen zusammengefasst. Die Kategorie **S** beinhaltet alle segmentierten komplexen Strukturen.

Tabelle 3: Strukturkategorien zur Analyse der Melodiestructur (Wermke, 2002; Wermke et al., 2007)

Struktur	Einfache Elemente
1B Einfachbogen	<p>Als Einfachbögen werden auf- und absteigende Melodiebögen definiert, die eine Länge von mindestens 300 ms und einen Hub von mindestens einem Ganztonschritt aufweisen.</p> 
KL Kurzlaute	<p>Eine Vokalisation mit einer Länge zwischen 45- 300 ms, wird als Kurzlaut kategorisiert.</p> 
Struktur	Komplexe Elemente
2B Doppelbogen	<p>Vokalisationen, deren Melodieverlauf aus zwei auf- und absteigenden Elementen besteht, werden als Doppelbogen kategorisiert. Bei dieser Kombination aus zwei Einzelbögen beträgt die Mindestlänge eines Einzelbogens 200 ms. Der Hub muss mindestens zwei Halbtonschritte pro Bogen aufweisen (Hubkriterium).</p> 

<p style="writing-mode: vertical-rl; transform: rotate(180deg);">3B Dreifachbogen</p>	<p>Vokalisationen mit Melodieverläufen, die die unter 2B angegebenen Kriterien erfüllen, jedoch aus drei auf- und absteigenden Elementen bestehen, wurden als Dreifachbogen klassifiziert.</p> 
<p style="writing-mode: vertical-rl; transform: rotate(180deg);">MB Mehrfachbögen</p>	<p>Zur Kategorie der Mehrfachbögen gehören alle Vokalisationen, die die unter 2B angegebenen Kriterien erfüllen, jedoch aus mehr als drei auf- und absteigenden Elementen bestehen.</p> 
<p style="writing-mode: vertical-rl; transform: rotate(180deg);">Segmentierungen</p>	<p>Als Segmentierungen werden Vokalisationen bezeichnet, in deren Grundfrequenzverlauf laryngeal bedingte Segmentierungspausen innerhalb einer Vokalisation auftreten, in denen keine Inspiration erfolgt (intentionaler Glottis-Verschluss). Eine vollständige Segmentierung liegt vor, wenn die Energiekurve um mindestens 9 dB absinkt und wieder ansteigt und eine vollständige Unterbrechung der Grundfrequenz und deren Harmonischen im Spektrogramm zu erkennen ist. Als unvollständig wird die Segmentierung gewertet, wenn im Spektrogramm noch Überreste der Harmonischen bzw. Rauschelemente zu erkennen sind, die nicht auf artikulatorische Aktivitäten zurückgeführt werden können. Je nach Ausprägung werden die segmentierten Vokalisationen weiter differenziert:</p>

<p style="writing-mode: vertical-rl; transform: rotate(180deg);">1S Einfache Segmentierung</p>	<p>Vokalisationen, innerhalb derer zwei einfache Elemente (1B oder KL) durch einen Intensitätsabfall und eine vollständige Unterbrechung der Grundfrequenz und deren Harmonischen getrennt sind werden als einfache Segmentierung klassifiziert.</p> 
<p style="writing-mode: vertical-rl; transform: rotate(180deg);">2S Zweifache Segmentierung</p>	<p>Zur Kategorie der Zweifachsegmentierung gehören Vokalisationen, deren Melodieverlauf zwei Unterbrechungen aufweisen. Als Elemente zwischen den Segmentierungen können sowohl Kurzlaute und Einfachbögen als auch Doppel- und Mehrfachbögen auftreten.</p>
<p style="writing-mode: vertical-rl; transform: rotate(180deg);">3S Dreifache Segmentierung</p>	<p>Hierzu zählen Vokalisationen, deren Melodieverlauf drei Unterbrechungen enthält. Als Elemente zwischen den Segmentierungen können sowohl Kurzlaute und Einfachbögen als auch Doppel- und Mehrfachbögen auftreten.</p> 
<p style="writing-mode: vertical-rl; transform: rotate(180deg);">MS Mehrfach-segmentierung</p>	<p>Alle Vokalisationen, deren Grundfrequenzverlauf mehr als drei Segmentierungspausen enthält zählen zur Kategorie der Mehrfachsegmentierung. Auch hier können als Elemente zwischen den Segmentierungen wieder sowohl Kurzlaute und Einfachbögen als auch Zwei- und Mehrfachbögen auftreten.</p>

KS Komplexe Segmentierung	<p>Für komplexe Segmentierungen gelten die gleichen Kriterien wie bei einer einfachen Segmentierung. Anders als bei einfachen Segmentierungen weist jedoch eines der beiden Elemente vor oder nach der Pause eine doppel-, drei- oder mehrbögeige Struktur auf.</p> 
Struktur	Weitere Elemente
??? Unklare Struktur	<p>In diese Kategorie fielen sämtliche Einzelvokalisationen, die aus verschiedenen Gründen (z.B. Überlagerungen, Hintergrundgeräusche, sehr starke F₀-Schwankungen) keiner der genannten Strukturen zuzuordnen sind. Sie werden in der weiteren Auswertung nicht berücksichtigt.</p>

4.5.4 Ermittlung des Melodiekomplexitätsgrades der Vokalisationen (MCI)

Analog zu Wermke et al. (2007) wurde in der vorliegenden Arbeit der *Melody Complexity Index (MCI)* pro Kind ermittelt als Maßzahl für das Verhältnis von einfachen zu komplexen Strukturen (Tabelle 3) für die kindbasierte Analyse. Er reflektiert den Anteil der Vokalisationen mit komplexen Melodiestructuren an der Vokalisationsgesamtzahl. Grundlage dafür war die Einteilung der Vokalisationen in einfache und komplexe Strukturen, welche auf Basis der Strukturkategorien (Kapitel 4.5.3) erfolgte. In vorliegender Arbeit wurden unter der Kategorie der einfachen Strukturen (eS) alle Einfachbögen und Kurzlaute zusammengefasst. Zur Kategorie der komplexen Strukturen (kpS) zählen alle Doppel- und Mehrfachbögen sowie sämtliche Segmentierungsformen. Der *MCI* definiert sich aus dem Quotienten der Anzahl von kpS zur Summe der Anzahl von kpS und eS und variiert zwischen 0 und 1 (Wermke et al., 2007):

$$MCI = \frac{kpS}{(kpS + eS)}$$

kpS: komplexe Strukturen (vgl. Tabelle 3); **eS:** einfache Strukturen (vgl. Tabelle 3)

Je größer der *MCI* ist, desto höher ist der Anteil komplexer Melodiestructuren im Vokalisationsrepertoire eines Kindes zu einem bestimmten Aufnahmezeitraum. Ziel der Analyse war es, die Entwicklung der Melodiekomplexität in Komfortvokalisationen zwischen dem 3. und 7. Monat längsschnittlich zu erfassen.

In der vorliegenden Arbeit wurde zunächst ein *MCI* pro Aufnahmetag pro Kind ermittelt. Auf Basis des arithmetischen Mittels dieser Werte konnte zunächst ein *MCI* für den gesamten Untersuchungszeitraum berechnet werden. Er spiegelt den Anteil komplexer Melodiestructuren im Vokalisationsrepertoire aller Kinder im Gesamtzeitraum wider. Anschließend wurden die *MCI*-Werte der einzelnen Aufnahmetage pro Kind gemittelt. Somit ergab sich der Anteil komplexer Strukturen pro Kind. Für die Darstellung des *MCI*s im Entwicklungsverlauf wurde zunächst pro Monat ein Durchschnittswert pro Kind berechnet. Das arithmetische Mittel dieser Werte spiegelt den Anteil komplexer Strukturen pro Monat wider und gibt somit Aufschluss darüber, wie sich der Anteil komplexer Strukturen in der Probandengruppe über den Untersuchungszeitraum verhält.

4.5.5 Quantitative Analyse – Erfassung akustischer Kenngrößen der Melodie

Neben der Analyse der Melodiestructuren wurden die Melodien aller Komfortvokalisationen interaktiv, mittels Cursor-Setzung, in *CDAP* vermessen. Dies diente der quantitativen Erfassung der Bogen- und Vokalisationslänge, der Erfassung der mittleren F_0 sowie des F_0 -Ranges (Frequenzmodulationsamplitude, F_0 -Hub) der Melodie der einzelnen Vokalisationen (siehe **Abbildung 4**). Die entsprechenden Werte werden automatisch in einer Excel-Tabelle abgelegt. Neben den einzelnen Melodiebögen wurden auch Segmentierungspausen sowie An- und Endwellen (kurze Ein- und Ausschwingungsvorgänge, < 150ms) vermessen. Letztere gingen nicht in die Berechnung der akustischen Eigenschaften mit ein. Die Berechnung der mittleren F_0 erfolgte auf Basis des geometrischen Mittels der Einzelvokalisationen. Der absolute F_0 -Range wurde durch die Differenz von F_{0max} und F_{0min} pro Laut gebildet. Der relative F_0 -Range berechnete sich aus dem Quotienten von F_{0max} und F_{0min} pro Laut. Beide Größen stellen ein Maß für den Dynamikbereich der mittleren F_0 , also den „Stimmumfang“ bei der Lautproduktion dar.

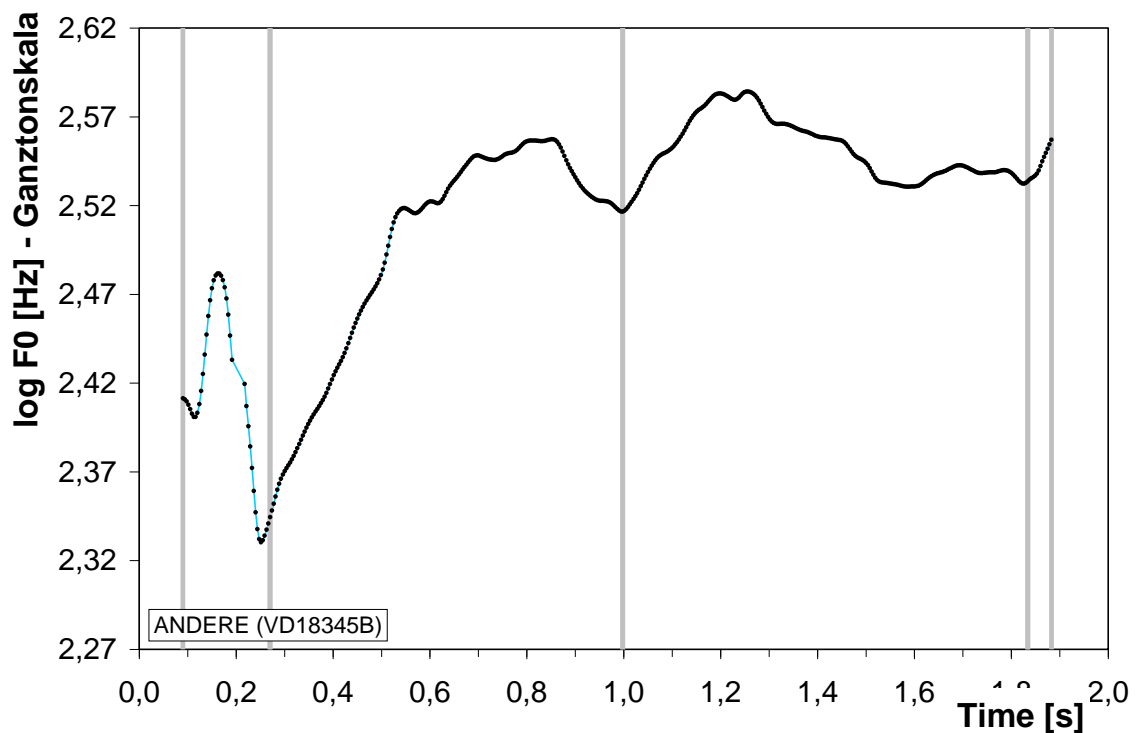


Abbildung 4: Vermessungsfenster im CDAP-Programm

Der blaue Graph zeigt die Melodiestructur. Die grauen vertikalen Linien spiegeln den Vermessungs-Cursor wider: Anlaut- Bogen 1- Bogen 2- Endlaut.

Die Cursor-Setzung erfolgte in der Regel an den Minima des Grundfrequenzverlaufs. War der Grundfrequenzverlauf (z.B. aufgrund von Störgeräuschen, Pressen) unterbrochen, erfolgte die Cursor-Setzung anhand des Intensitätsverlaufs. Dadurch konnte auch bei unvollständigem Grundfrequenzverlauf zumindest die Vokalisationslänge ermittelt werden. Diese basiert auf der Summe der einzelnen Bogen- und Pausen-Längen innerhalb einer Vokalisation. In die Berechnung der Vokalisationslänge gingen auch Kurzlaute (45-300ms) ein. Diese gehören zum Grundrepertoire der untersuchten vorsprachlichen Phase und sollten demnach nicht ausgeschlossen werden (Oller, 2000).

4.5.6 Analyse artikulatorischer Fähigkeiten

Neben einer eher larynx-fokussierten Analyse der stimmphysiologischen Parameter und der Melodiekomplexität, wurden parallel artikulatorische Aktivitäten auf der supralaryngealen Ebene analysiert. Um mögliche Beziehungen zwischen der melodischen und artikulatorischen Entwicklung aufzudecken, erfolgte die Analyse von supralaryngealen Aktivitäten nicht nur auf Vokalisationsebene, sondern auch auf Ebene der Melodiestruktur. Dies erforderte mehrere Analyseschritte, welche nachfolgend erläutert werden. Neben der Darstellung der Frequenzverläufe sowie der artikulatorischen Elemente in Schmal- und Breitbandspektrogrammen mittels *Praat-Software* erfolgte die Kodierung der artikulatorischen Elemente parallel in einem geeigneten excel-basierten Analysemodul in *CDAP* (Anmerkungsfenster).

4.5.6.1 Erfassung artikulatorischer Elemente pro Vokalisation

Die Identifizierung und Lokalisierung artikulatorischer Aktivitäten innerhalb einer Äußerung erfolgte mittels auditiver Kontrolle sowie einer visuellen Analyse mit Hilfe der Ausgabe des zugehörigen Breitband- und Schmalbandspektrogramms im Ausgabefenster der *Praat-Software* (siehe **Abbildung 5** und **Abbildung 7**). Interaktiv konnte im Ausgabefenster des jeweiligen Spektrogramms auch der Melodieverlauf abgebildet werden. Im Unterschied zum Schmalbandspektrogramm hat das Breitbandspektrogramm eine hohe zeitliche Auflösung und eignet sich daher v.a. zur Identifizierung und Lokalisierung von konsonantischen Elementen (z.B. Plosiven).

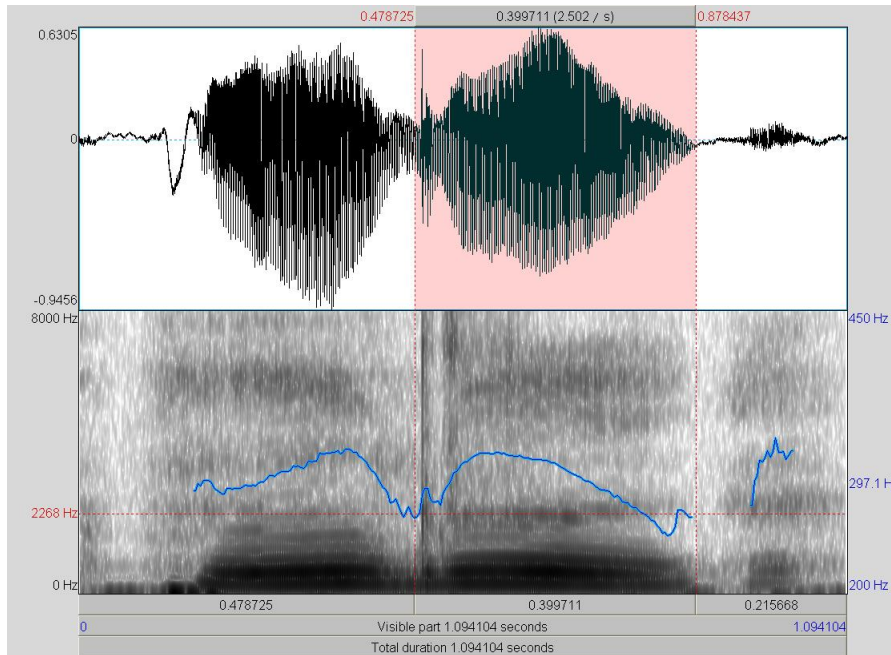


Abbildung 5: Breitbandspektrum (unteres Fenster) mit Amplitudenverlauf (oberes Fenster) im Ausgabefenster der *PRAAT*-Software.

Der blaue Graph zeigt den Melodieverlauf (Doppelbogen mit Artikulation am Rand des 2. Bogens) gegen die Zeit. Auf der y-Achse ist der Frequenzbereich der Breitbanddarstellung (Hz) abgebildet (rote Ziffer). Auf der z-Achse lässt sich der Frequenzbereich (Hz) der Melodiekontur ablesen (blaue Ziffer). Auf der x-Achse ist die Zeit (s) dargestellt. Der rot hinterlegte Bereich kennzeichnet jeweils den Cursorbereich und erleichtert die auditive Kontrolle.

Die Analyse von Artikulationsbewegungen erfolgte ausschließlich bei Vokalisationen, die eine Mindestlänge von 300 ms aufwiesen. Es stellte sich heraus, dass die auditive Kontrolle spektraler Eigenschaften bei Vokalisationslängen <300 ms (Kurzlaute) kaum möglich ist. Die **Abbildung 6** zeigt das von der Autorin entworfene und verwendete Kodierungsschema von Artikulationsbewegungen innerhalb einer Vokalisation. Die nachfolgenden Kapitel erklären das Schema genauer.

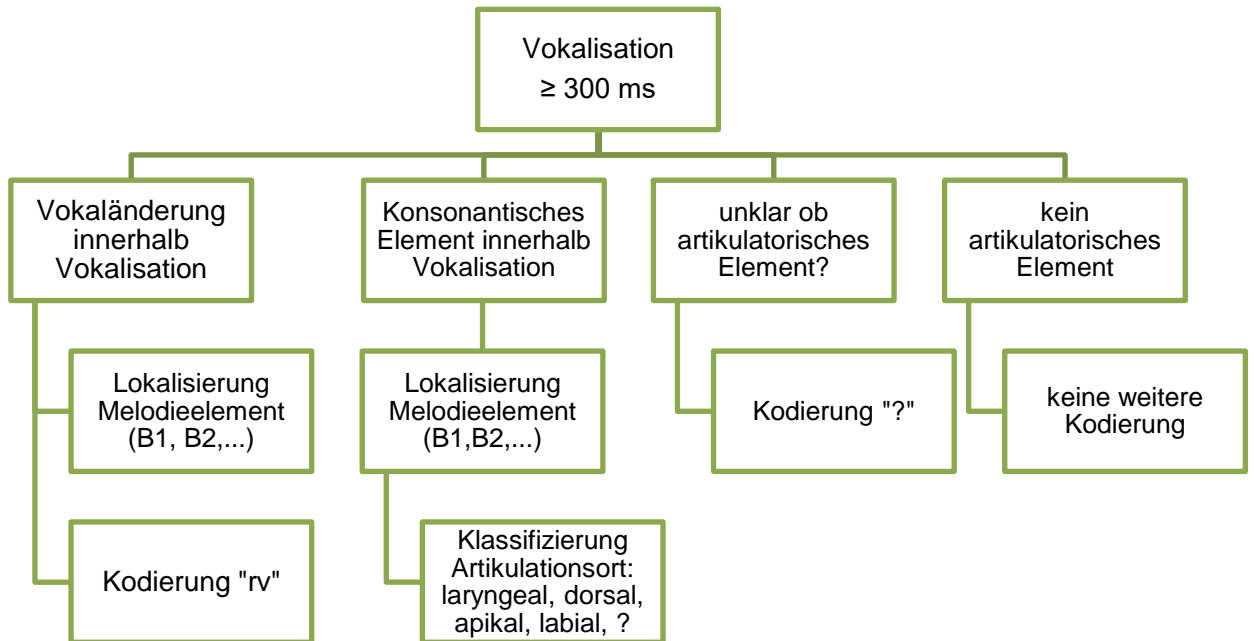


Abbildung 6: Kodierungsschema von Artikulationsbewegungen innerhalb einer Vokalisation.

4.5.6.2 Analyse von Resonanzveränderungen

Wenn innerhalb einer Vokalisation eine Resonanzveränderung (Vokaländerung) stattfand, wurde diese hinter dem jeweiligen Melodiebogen mit der Abkürzung „rv“ kodiert. Die Kodierung „rv“ wurde vergeben, wenn eine deutliche Energieveränderung der Harmonischen in diesem Bereich erkennbar war, ein Anstieg oder Abfall des ersten Formanten (F1) erkennbar und die Vokaländerung eindeutig auditiv wahrnehmbar war (siehe **Abbildung 7**).

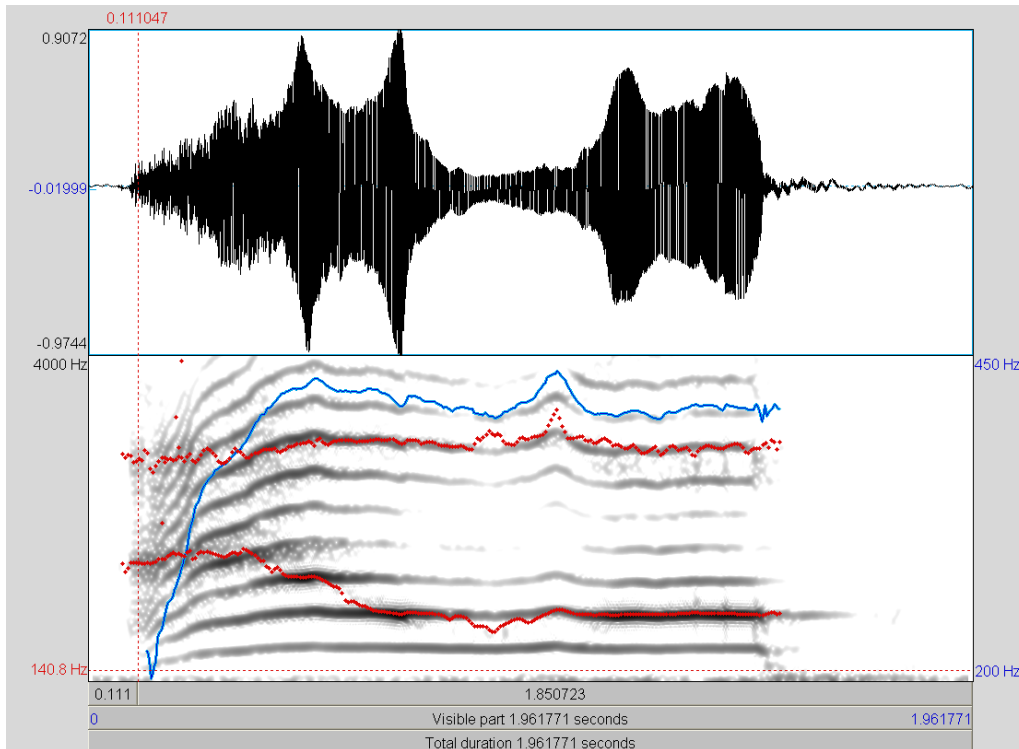


Abbildung 7: Darstellung einer Resonanzveränderung (rv) innerhalb einer Vokalisation im Schmalbandspektrum im Ausgabefenster der *PRAAT*-Software.

Der blaue Graph zeigt den Melodieverlauf gegen die Zeit. Auf der y-Achse ist der Frequenzbereich (Hz) der Schmalbanddarstellung abgebildet. Auf der z-Achse lässt sich der Frequenzbereich (Hz) der Melodiekontur ablesen. Auf der x-Achse ist die Zeit (s) dargestellt. Die beiden roten Graphen spiegeln den mittels LPC ermittelten ersten (F1) und zweiten (F2) Formanten wider. Der Wechsel von Hell zu Dunkel im Bereich der oberen Harmonischen spiegelt eine Energieveränderung wider.

4.5.6.3 Analyse konsonanten-ähnlicher Elemente

Ein weiterer Analyseschritt bestand darin, auftretende konsonanten-ähnliche Elemente genauer zu beschreiben. Wurde sowohl auditiv als auch anhand spektraler Eigenschaften ein konsonanten-ähnliches artikulatorisches Element innerhalb einer Äußerung identifiziert, wurde, für eventuelle spätere Analysen, zusätzlich ermittelt, zu welchem Melodieelement (Bogen) die Artikulation in Beziehung stand (silbenbasierte Analyse).

Zudem erfolgte auditiv und spektral eine Klassifizierung des Artikulationsortes (siehe **Abbildung 8**). **Tabelle 4** gibt Aufschluss über die Kodierungskriterien der Artikulationsbewegungen innerhalb einer Melodiestructur. Dabei handelte es sich nicht um Kriterien für phonetisch korrekt realisierte Konsonanten der deutschen Sprache,

sondern lediglich um die Beschreibung des Artikulationsortes bzw. der Zungenposition im Vokaltrakt (laryngeal, dorsal, apikal, labial). Aus Gründen der Vereinfachung wurde pro Vokalisation entweder „rv“ **oder** das Auftreten konsonantischer Elemente kodiert. Die objektiven Kriterien einer Resonanzveränderung wären beim Auftreten zusätzlicher konsonantischer Elemente aufgrund koartikulatorischer Effekte nicht mehr eindeutig zu bestimmen gewesen. Des Weiteren wurde die Produktion konsonanten-ähnlicher Elemente als höhere Leistung interpretiert und erfasst.

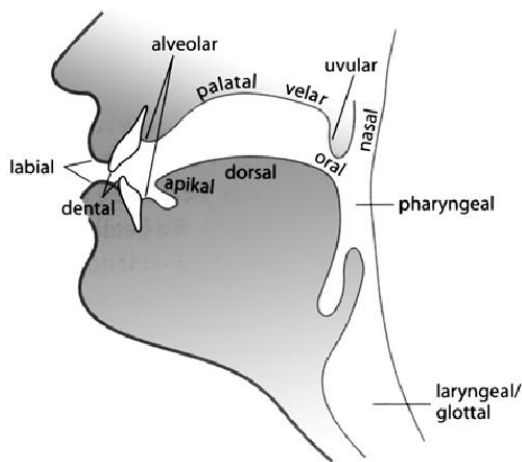
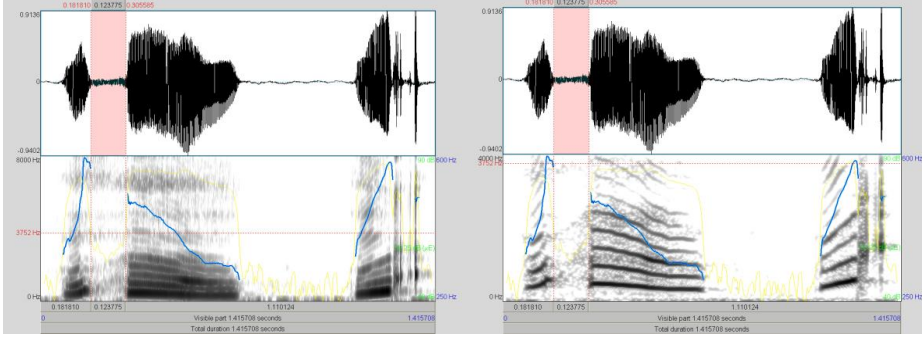


Abbildung 8: Artikulationsbereiche und -organe des Mundraums (Weinrich & Zehner, 2017), S.7.

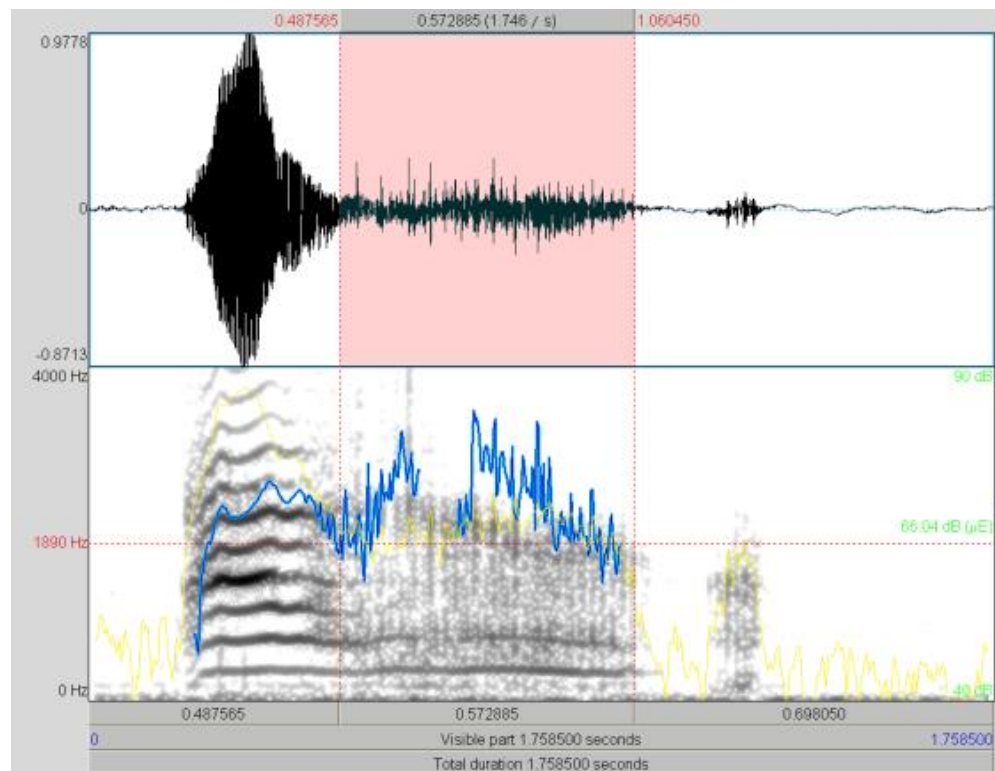
Tabelle 4: Kodierung der Artikulationsbewegungen

Ebene	Laryngeale Artikulation
Laryngeal	<p>Als laryngeal wird lediglich die Produktion des Engelautes [h] in medialer Position bezeichnet. Die Enge entsteht dabei direkt an der Glottis, also auf laryngealer Ebene. Im Spektrogramm ähnelt die Struktur sehr einer Segmentierungspause, da der Grundfrequenzverlauf meistens an Stelle der Artikulationsbewegung unterbrochen ist. Auditiv ist die Artikulation jedoch eindeutig wahrzunehmen. Initial, also am Äußerungsbeginn, wird der Engelaute [h] nicht als artikulatorisches Element behandelt, da er an dieser Stelle eher als ein unwillkürliches „Einschwingen“ gewertet werden muss, welches nicht als bewusste artikulatorische Aktivität zu interpretieren ist.</p>
	<div style="text-align: center;">  <p data-bbox="491 1128 638 1160">B1(laryngeal)</p> <p data-bbox="411 1187 1398 1303">Laryngeale Artikulation (roter Cursorbereich: Unterbrechung der Grundfrequenz und der Harmonischen) im Breitbandspektrogramm (links) und Schmalbandspektrogramm (rechts) mit anschließendem Inspirationslaut.</p> </div>

Dorsale Artikulation

Als dorsal werden sämtliche Artikulationsbewegungen klassifiziert, die unter Beteiligung des hinteren (dorsalen) Bereichs der Zunge [r, g, ng, k] oder durch eine Verengung des Pharynx [ch] ausgeführt werden. Der dorsale Reibelaut [r] ist im Spektrogramm meistens durch einen sehr „gezackten“ F_0 -Verlauf zu erkennen. Dorsale Verschlusslaute hingegen sind durch kurze vertikale Linien im Spektrogramm zu erkennen. Auch hier entsteht oft der Eindruck von Segmentierungspausen. Allerdings handelt es sich hierbei nur um die Zeit der Verschlusslösung.

supralaryngeal



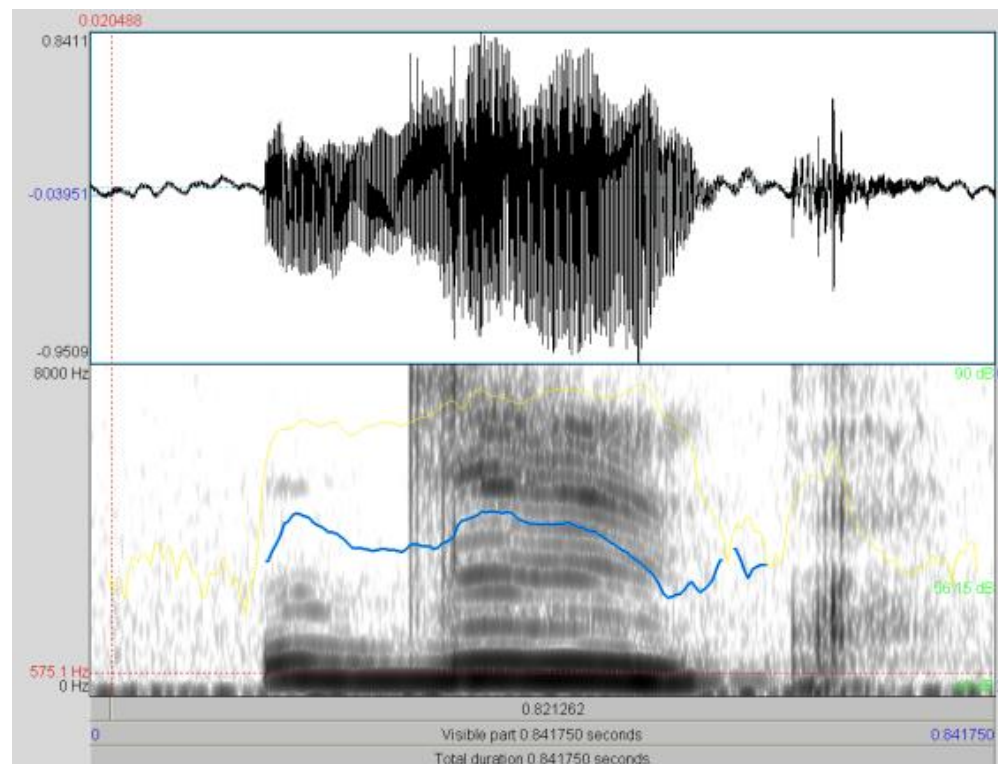
B1 – B2(dorsal)

Dorsale Artikulation –r- (roter Cursorbereich: starke Turbulenzen der Grundfrequenz) auf dem zweiten Melodiebogen im Schmalbandspektrogramm.

Apikale Artikulation

Als apikale Artikulation werden sämtliche Artikulationsbewegungen bezeichnet, die mit dem vorderen Anteil der Zunge (lat.: apex) ausgeführt werden. Das kann sowohl in Form von Nasalen, z.B. [n], aber auch in Form von Verschlusslauten, Lateralen und Frikativen, z.B. [d, t, s, sch, l] der Fall sein. Nasale sind meistens durch einen Energieabfall in den oberen Harmonischen zu erkennen, während die apikalen Verschlusslaute ähnlich wie die dorsalen Verschlusslaute durch eine dünne vertikale Linie im Spektrogramm zu erkennen sind.

supralaryngeal



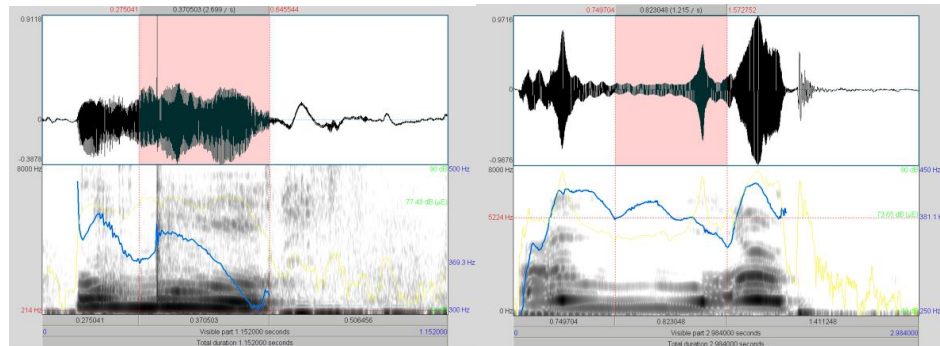
A – B1(apikal) - E

Apikale Artikulation –nd- (Abfall der Energie und Auslöschung der oberen Harmonischen im Bereich der nasalen Artikulation; vertikale Linie im Bereich des Verschlusslösung der plosiven Artikulation) im Breitbandspektrogramm.

Labiale Artikulation

Als labiale Artikulation werden sämtliche artikulatorischen Elemente bezeichnet, welche durch Beteiligung der Lippen (bzw. nur der Ober- oder Unterlippe) produziert werden. Hierzu zählen der Nasal [m], Verschlusslaute [p, b], Reibelauten [f, w] und sog. *raspberries*.

supralaryngeal



A – B1 – B2 (labial) – E

B1 (labial) – B2 (labial) – B3 (labial)

Labiale Artikulation –b- (links) und –w- (rechts) im Breitbandspektrum

links: vertikale Linie im Bereich der Verschlusslösung am Rand des zweiten Bogens;

rechts: Energieabfall und Auslöschung der oberen Harmonischen im Bereich des Reibelautes auf dem zweiten Bogen und den Rändern des ersten und dritten Bogens.

unklare Artikulation

Als unklare Artikulation werden Artikulationsbewegungen bezeichnet, die weder auditiv noch anhand spektraler Eigenschaften einem Artikulationsorgan/-ort zugeordnet werden können. Dies ist beispielsweise beim Auftreten von Störgeräuschen oder schlechter Aufnahmequalität der Fall. Da die Artikulationsbewegungen anfänglich noch sehr verwaschen sind und nicht den erwachsenensprachlichen Konsonanten entsprechen, ist es in einigen Fällen nicht möglich sie eindeutig einer Artikulationsebene zuzuordnen. Daher wird nur die Position innerhalb der Melodistruktur bestimmt und für die Artikulationsebene ein Fragezeichen (?) vergeben. Falls nicht festgestellt werden konnte, ob es sich um eine Artikulationsbewegung oder ein Störgeräusch handelt, wurde lediglich ein Fragezeichen vergeben.

Bsp.:

B1 - B2(?) - B3(?)

4.6 Statistische Analyse

Die statistische Datenanalyse erfolgte mithilfe von *Microsoft-Excel* (Version 2010) und *SPSS Statistics* (Version 26). Zunächst wurden in einer explorativen Datenanalyse die Häufigkeiten sowie die Lage- und Streuungsmaße aller Analyseparameter auf Vokalisationsebene und Kindebene dargestellt. Anschließend erfolgte mit dem K-S-Test für alle Analyseparameter eine Prüfung auf Normalverteilung. Lag keine Normalverteilung vor, wurden für die entsprechenden kindbasierten statistischen Analysen nichtparametrische Tests verwendet.

5. Ergebnisse

5.1 Charakterisierung akustischer Kenngrößen von Komfort-Vokalisationen im Alter von 3-7 Monaten

Die Analysen der akustischen Kenngrößen erfolgten, wie in Kapitel 4 beschrieben und begründet, sowohl lautbasiert als auch kindbasiert.

5.1.1 Mittlere Grundfrequenz und F_0 -Range (Hub)

5.1.1.1 Lautbasierte Analyse der mittleren Grundfrequenz und des F_0 -Ranges

Die Charakterisierung der mittleren F_0 sowie die des F_0 -Ranges wurden hier auf Basis aller Einzel-Vokalisationen, ohne vorher durchgeführte Kind-Mittelung, durchgeführt. In **Tabelle 5** und **Tabelle 6** sind die statistischen Kenngrößen (Mittelwert, Standardabweichung, Median, Perzentile, Interquartil-Bereich, Minimum, Maximum) für die mittlere Grundfrequenz sowie für den F_0 -Range (absolut und relativ) für das gesamte Datenset aufgeführt.

Für die mittlere F_0 ergab sich ein Median von 323,1 Hz. Der Interquartil-Bereich erstreckt sich dabei über 96 Hz, also etwa 5 Halbtöne. Der F_0 -Mittelwert beträgt 360,6 Hz. Die Standardabweichung von 168,7 Hz zeigt, dass die Streuung der Daten sehr hoch ist. Die Untersuchung der statistischen Verteilung der F_0 -Werte ergab keine hinreichende Normalverteilung (Kolmogorov-Smirnov-Test), siehe Histogramm **Abbildung 9**.

Der Boxplot in **Abbildung 9** gibt Aufschluss über die zentrale Tendenz der mittleren F_0 -Daten. Der Großteil der Daten ist zwar relativ gleichmäßig verteilt, dennoch zeigt sich eine gewisse Rechtschiefe aufgrund der Ausreißer im oberen Frequenzbereich. Hierbei ist anzumerken, dass es sich dabei um in diesem Altersbereich typische Lautproduktionen handelt.

Tabelle 5: Deskriptive Statistik der mittleren F_0 im Gesamtzeitraum

N = 5475		Statistik	
Mittlere F_0 [Hz]	Mittelwert	360,6	
	Standardabweichung	168,7	
	Median	323,1	
	Perzentile	25.	281,6
		75.	377,7
	Interquartilbereich	96,0	
	Minimum	120,9	
Maximum	1923,6		

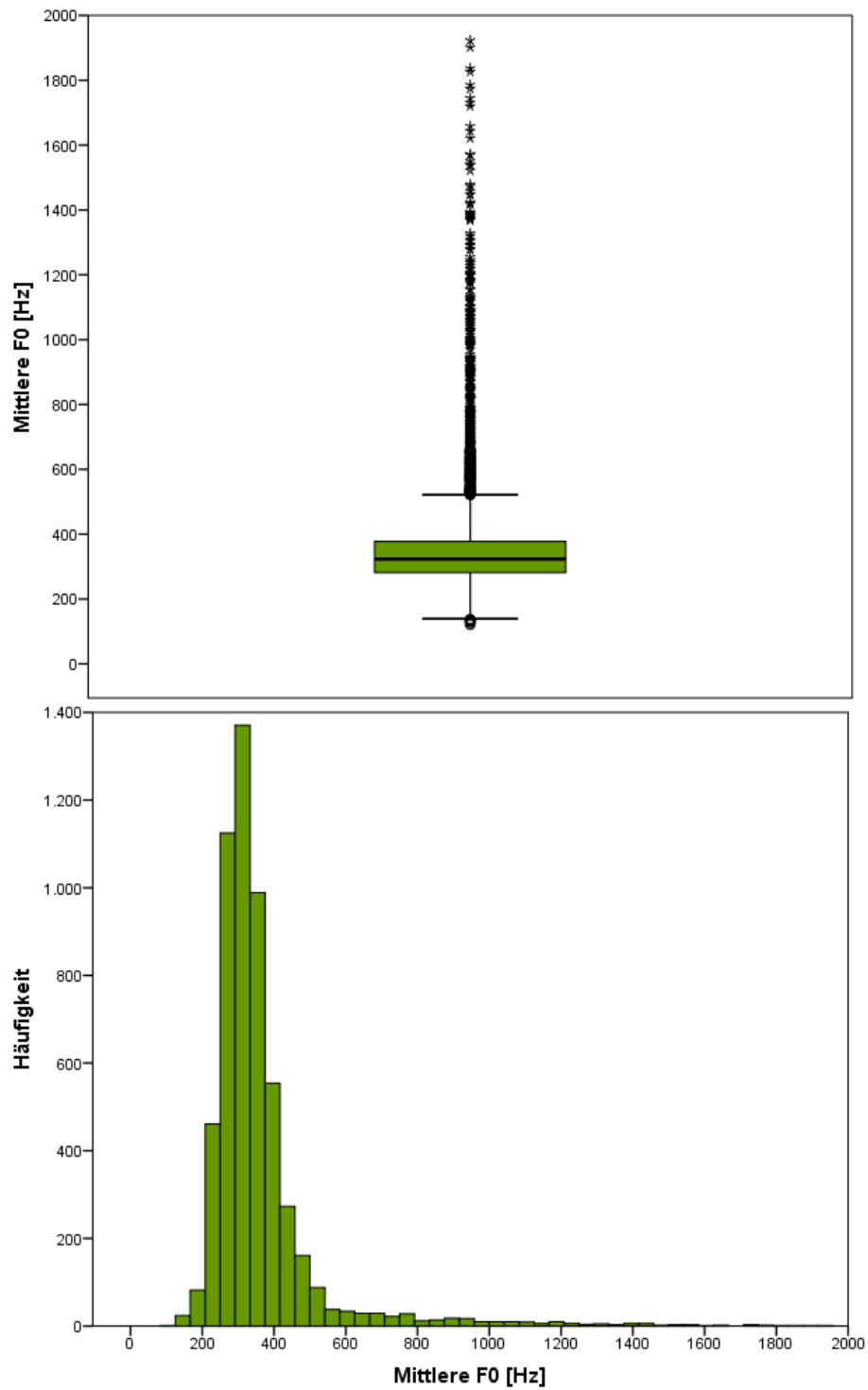


Abbildung 9: Boxplot (oben) und Histogramm (unten) der mittleren F_0 -Daten im Gesamtzeitraum

Tabelle 6 zeigt die Verteilungseigenschaften des absoluten und relativen F_0 -Ranges im Gesamtzeitraum. Die Angabe des relativen F_0 -Ranges erfolgte ergänzend, da sie der Wahrnehmung von Frequenzunterschieden des menschlichen Gehörs besser gerecht wird. Boxplot und Histogramm in **Abbildung 10** zeigen, ähnlich wie bei der mittleren F_0 , eine gewisse Rechtsschiefe.

Tabelle 6: Deskriptive Statistik von absolutem und relativem F_0 -Hub im Gesamtzeitraum

N= 5475	Mittelwert	SD	Median	Perzentile		Interquartil	Minimum	Maximum
Absoluter F_0-Hub [Hz]	174,7	196,1	124,1	25.	75,7	119,2	4,2	2831,2
				75.	194,9			
Relativer F_0-Hub	1,7	1,0	1,5	25.	1,3	0,6	1,0	18,7
				75.	1,8			

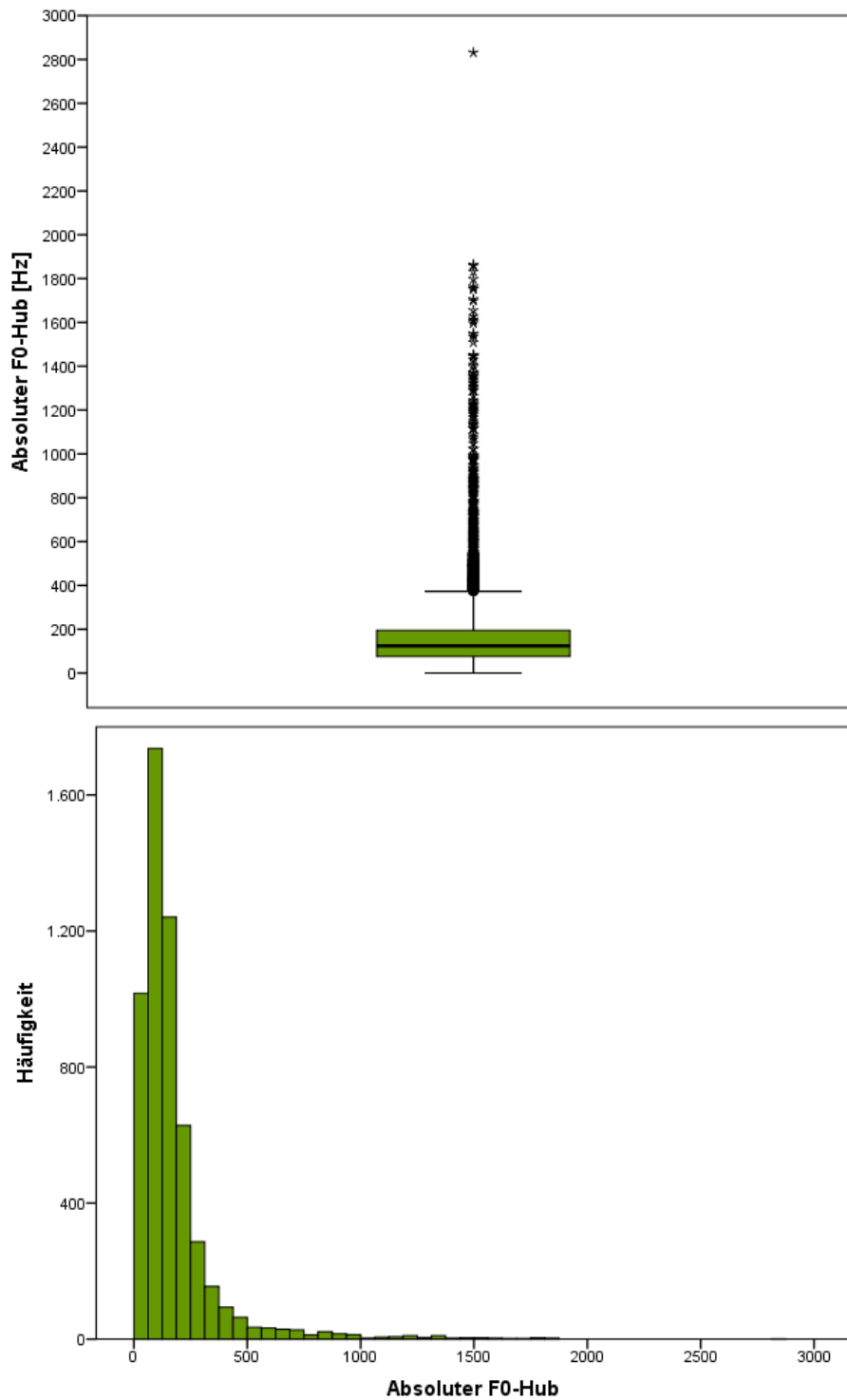


Abbildung 10: Boxplot (oben) und Histogramm (unten) des absoluten F_0 -Hubes im Gesamtzeitraum

Im Folgenden werden die Ergebnisse der Analysen der Mittleren F_0 sowie des F_0 -Ranges in monatlichen Zeitabständen (3. Monat - 7. Monat) berichtet.

Tabelle 7 und **Tabelle 8** enthalten die Mittelwerte, Standardabweichungen, Mediane, Perzentile sowie Minimum und Maximum für die mittlere F_0 und für den absoluten F_0 -Range pro Monat.

Tabelle 7: Deskriptive Statistik der mittleren F_0 pro Monat

Alter in Monaten		3 (n=863)	4 (n=1828)	5 (n=1237)	6 (n=1153)	7 (n=394)	
Mittlere F_0 [Hz]	Mittelwert	324,9	318,3	369,2	416,8	443,8	
	SD	82,6	76,7	183,9	243,4	215,3	
	Median	313,3	310,1	324,0	338,3	390,8	
	Perzentile	25.	276,5	273,0	285,2	287,9	322,7
		75.	357,7	353,9	371,1	447,4	472,7
	Interquartilbereich	81,2	80,8	85,9	159,5	150,0	
	Minimum	128,3	120,9	148,0	137,6	177,5	
Maximum	1072,83	1540,8	1643,3	1923,6	1473,0		

Der Boxplot in **Abbildung 11** verdeutlicht, dass sich die mittlere F_0 im 3. und 4. Monat relativ stabil verhält. Ab dem 5. Monat kommt es zu einer Zunahme der Variation durch hochfrequenter Vokalisationen. Dieser Trend bleibt in Monat 6 und Monat 7 bestehen, wobei sich die Interquartilbereiche vergrößern. Die Spannweite ist im 6. Monat größer als zu anderen Zeitpunkten.

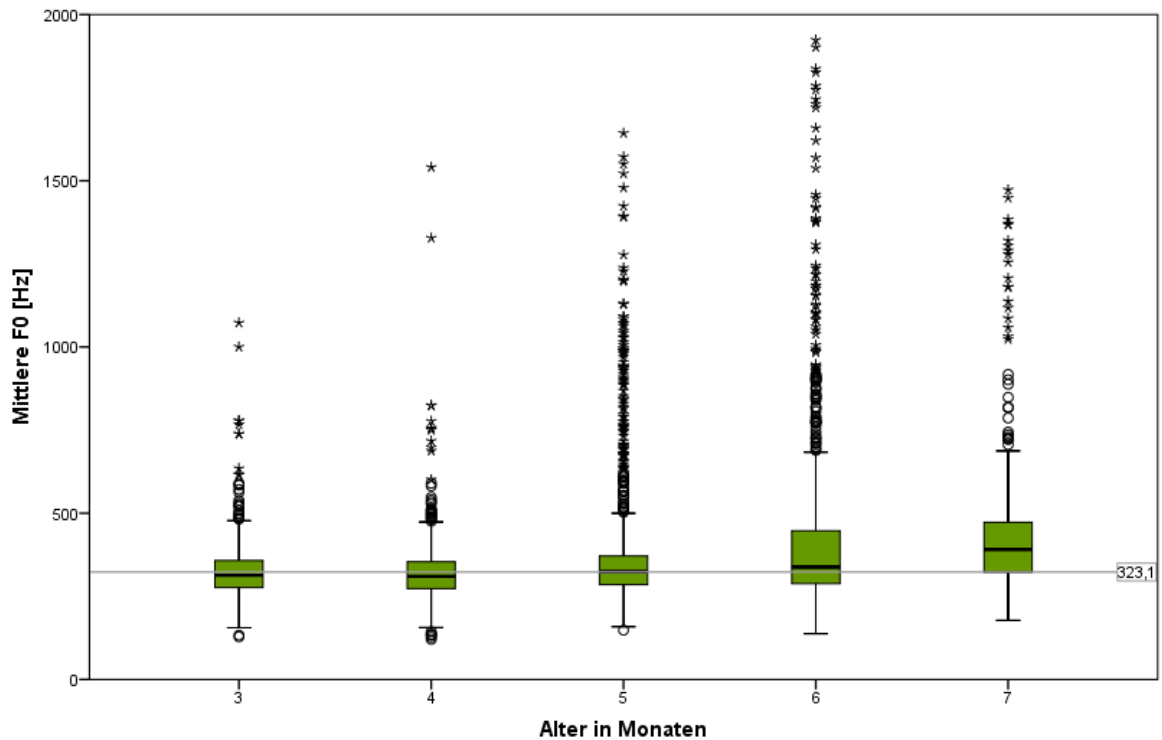


Abbildung 11: Boxplot der mittleren F_0 für die einzelnen Monate im Untersuchungszeitraum mit eingezeichnetem Gruppenmedian (graue Linie)

Im Unterschied zur mittleren F_0 sind die Werte des absoluten F_0 -Ranges im Untersuchungszeitraum relativ stabil (Abbildung 12). Dennoch zeigt sich auch hier ab dem 5. Monat eine höhere Variabilität. Während der Median im 3. und 4. Monat noch unter dem Gruppenmedian von 124,1 Hz liegt, steigt er zwischen dem 5. und 7. Monat von 203,8 Hz auf 263,7 Hz an. Ebenso vergrößert sich der Interquartil-Abstand ab dem 5. Monat um mehr als die Hälfte. Auch die Spannweite des F_0 Hubes ist ab dem 5. Monat größer als zu früheren Zeitpunkten.

Tabelle 8: Deskriptive Statistik des absoluten F₀-Hubes pro Monat

Alter in Monaten		3	4	5	6	7	
		(n=863)	(n=1828)	(n=1237)	(n=1153)	(n=394)	
Absoluter F ₀ -Hub [Hz]	Mittelwert	138,3	136,8	185,3	234,2	222,1	
	SD	135,2	121,3	203,8	264,1	263,7	
	Median	105,7	111,6	133,3	156,0	156,3	
	Perzentile	25.	63,6	72,5	77,1	86,9	93,3
		75.	170,3	163,4	207,1	263,4	248,2
	Interquartilbereich	106,8	90,9	130,0	176,5	154,9	
	Minimum	11,1	0,0	4,2	5,2	11,3	
Maximum	1535,4	1367,2	1797,7	1863,6	2831,2		

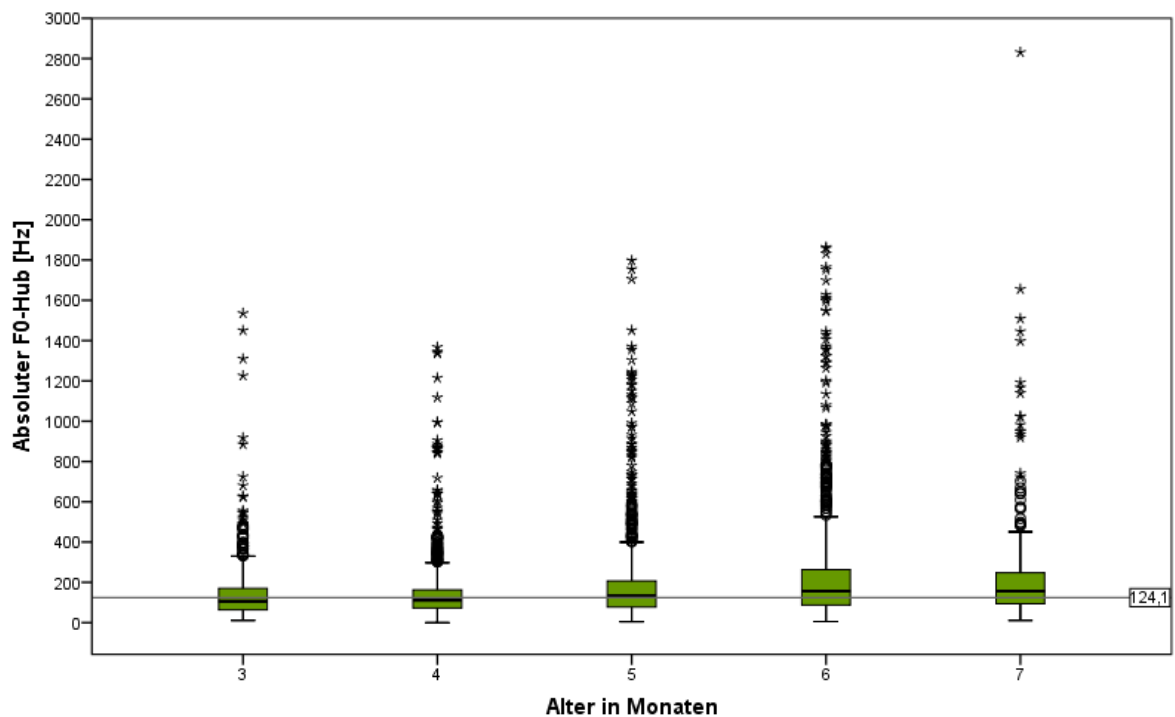
Abbildung 12: Boxplot für den absoluten F₀-Hub innerhalb einer Vokalisation pro Monat mit eingezeichnetem Gruppenmedian (graue Linie)

Tabelle 9 liefert die Ergebnisse der lautbasierten Analyse der mittleren F_0 pro Kind. Der Boxplot in **Abbildung 13** zeigt, dass sich die Mediane der einzelnen Kinder relativ eng um den Gruppen-Median bewegen. Auch der Interquartil-Bereich scheint relativ ähnlich für die einzelnen Kinder zu sein. Die Breite der Verteilungseigenschaften, also die Spannbreite, unterscheidet sich jedoch teilweise stark zwischen den einzelnen Kindern. Während die Minima der mittleren F_0 innerhalb der Gruppe relativ ähnlich sind, liegen die Maxima einiger Kinder mehr als doppelt so hoch wie die von anderen Kindern. So produzieren beispielsweise die Kinder AC, AD und VH im Vergleich zum Rest der Kinder keine Laute mit einer mittleren F_0 über 1000 Hz (siehe **Tabelle 9**).

Tabelle 9: Deskriptive Statistik der mittleren F_0 pro Kind im Untersuchungszeitraum

Mittlere F_0 [Hz] N= 5475		Mittelwert	SD	Median	Perzentile	Interquartil	Minimum	Maximum	
VF	♂	278,7	80,0	264,4	25.	242,9	45,4	136,0	1197,8
					75.	288,3			
AC	♂	293,8	48,4	296,2	25.	272,5	46,4	127,4	590,3
					75.	318,9			
AD	♀	321,6	65,4	314,4	25.	285,5	66,7	133,0	824,4
					75.	352,2			
VE	♂	330,6	120,4	302,7	25.	268,4	88,4	155,6	1244,6
					75.	356,7			
VI	♂	333,1	149,9	301,5	25.	268,7	62,2	120,9	1658,7
					75.	330,8			
VH	♀	359,1	51,3	355,3	25.	325,2	65,7	208,5	584,9
					75.	390,9			
VD	♀	360,7	175,7	303,7	25.	272,9	84,6	128,3	1572,2
					75.	357,5			
AK	♀	374,1	153,9	342,0	25.	309,4	68,9	138,9	1387,5
					75.	378,3			
VB	♀	413,5	262,7	332,9	25.	293,8	108,8	159,6	1836,9
					75.	402,7			
VL	♀	420,4	285,7	328,4	25.	296,6	87,4	148,0	1923,6
					75.	384,0			
VG	♂	421,2	149,4	380,0	25.	346,9	103,3	137,6	1448,4
					75.	450,3			

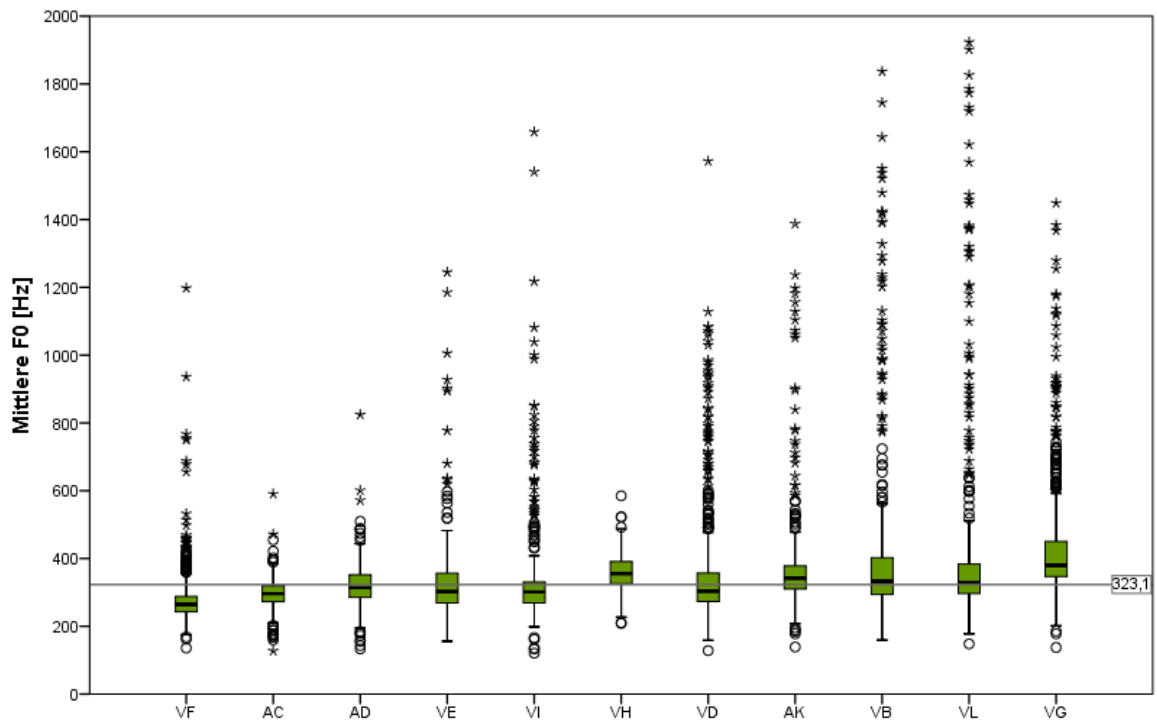
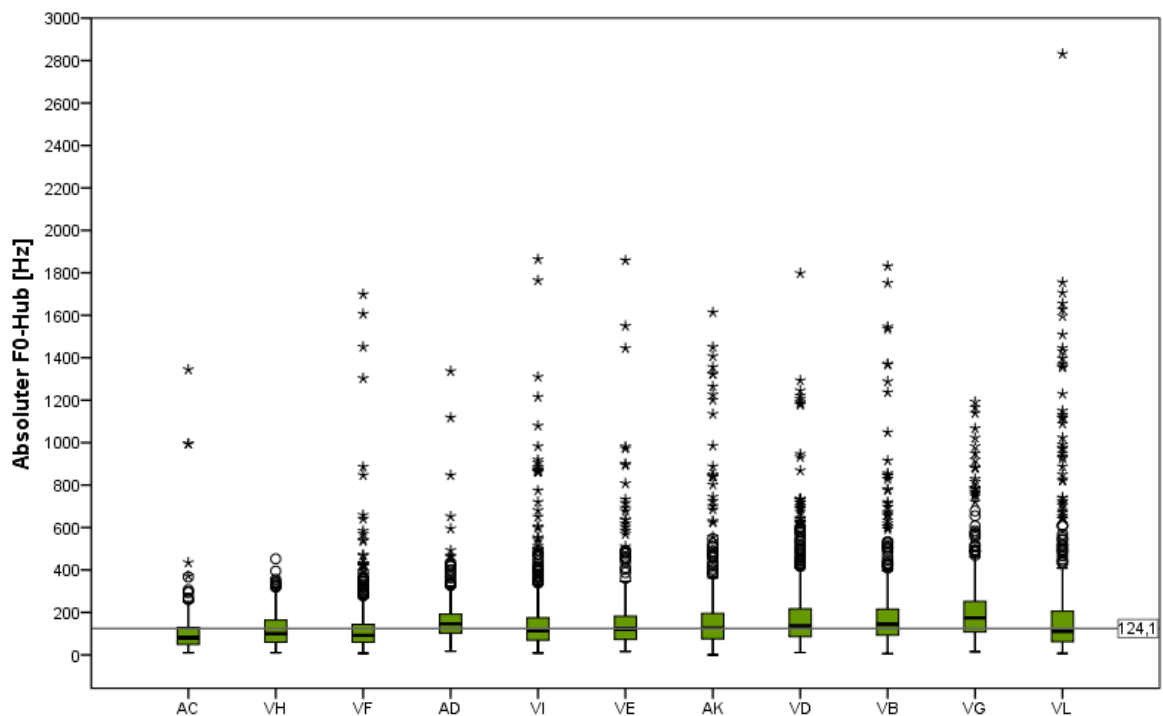


Abbildung 13: Boxplot der mittleren F_0 für die einzelnen Kinder mit eingezeichnetem Gruppenmedian

In **Tabelle 10** ist die deskriptive Statistik für den absoluten F_0 -Range pro Kind aufgelistet. Der dazugehörige Boxplot in **Abbildung 14** zeigt, dass der Interquartil-Abstand und die Verteilung der Daten relativ ähnlich für die meisten Kinder sind. Allerdings weisen auch hier die beiden Kinder AC und VH kaum bzw. gar keine Extremwerte auf. Bei dem Laut-Maximum von 2831,2 Hz bei dem Kind VL handelt es sich um einen echten extrem hochfrequenten Freuden-Quieker, welcher alle Laut-Auswahlkriterien erfüllte und daher aus Gründen der Vollständigkeit mit aufgeführt wurde (**Abbildung 14**).

Tabelle 10: Deskriptive Statistik für den absoluten F₀-Hub pro Kind

Absoluter F ₀ -Hub [Hz] N= 5475		Mittelwert	SD	Median	Perzentile		Interquartil	Minimum	Maximum
					25.	75.			
AC	♂	106,4	120,9	80,5	25.	49,5	80,6	10,2	1344,1
					75.	130,1			
VH	♀	119,4	78,1	100,3	25.	59,4	104,5	9,9	452,4
					75.	163,9			
VF	♂	127,8	150,3	91,7	25.	59,8	83,8	7,5	1698,7
					75.	143,6			
AD	♀	167,0	123,1	146,1	25.	102,8	89,1	17,3	1336,3
					75.	191,9			
VI	♂	168,5	202,9	111,8	25.	69,7	106,2	8,3	1863,6
					75.	175,9			
VE	♂	170,0	195,3	122,4	25.	72,9	110,9	15,7	1858,4
					75.	183,8			
AK	♂	191,3	230,7	127,0	25.	75,2	120,1	0,0	1613,2
					75.	195,3			
VD	♀	199,1	208,8	135,8	25.	85,4	132,2	10,9	1797,7
					75.	217,5			
VB	♀	202,4	230,3	144,3	25.	92,6	123,12	5,2	1831,5
					75.	215,7			
VG	♂	206,1	152,4	174,7	25.	108,2	144,0	15,5	1191,7
					75.	252,3			
VL	♀	219,0	320,3	110,1	25.	62,2	143,6	6,2	2831,2
					75.	205,8			

**Abbildung 14:** Boxplot für den absoluten F₀-Hub [Hz] pro Kind mit eingezeichnetem Gruppenmedian (graue Linie)

5.1.1.2 Kindbasierte Analyse der mittleren Grundfrequenz und des F_0 -Ranges

Ein weiterer Ansatz bei der Datenanalyse bestand darin, die akustischen Kenngrößen der Vokalisationen pro Monat auf Basis der einzelnen Kinder zu mitteln (siehe **Tabelle 11**). Der Boxplot in **Abbildung 15** zeigt einen deutlichen Anstieg der Mediane der mittleren F_0 mit einem Peak im 6. Monat. Eine Rangvarianzanalyse durch den Friedman-Test ($p=0,26$) ergab jedoch keine signifikanten Unterschiede der mittleren F_0 im Untersuchungszeitraum. Dennoch scheinen sich die Veränderungen des Lautrepertoires zwischen dem 3. und 7. Monat auch in der kindbasierten Analyse widerzuspiegeln. Ebenso verdeutlicht der Boxplot die interindividuelle Variabilität bei der Produktion von Komfortvokalisationen, welche in vorliegender Probandengruppe ab dem 5. Monat größer wird (Expansion-Stage).

Tabelle 11: Deskriptive Statistik der mittleren F_0 pro Monat (kindbasierte Mittelung)

Alter in Monaten		3 (n=11)	4 (n=11)	5 (n=10)	6 (n=10)	7 (n=6)	
Mittlere F_0 [Hz]	Mittelwert	329,7	321,4	379,5	430,3	390,5	
	SD	39,5	29,9	147,5	128,2	93,6	
	Median	326	317,7	341,4	410	377,7	
	Perzentile	25.	306	300	282,3	341,3	308,7
		75.	357	357,3	396,3	488,8	467,7
	Interquartilbereich	51	57,3	114	147,5	159	
	Minimum	260,5	366,4	772,38	268,3	285,5	
Maximum	408,6	1540,8	1643,3	706,8	543,4		

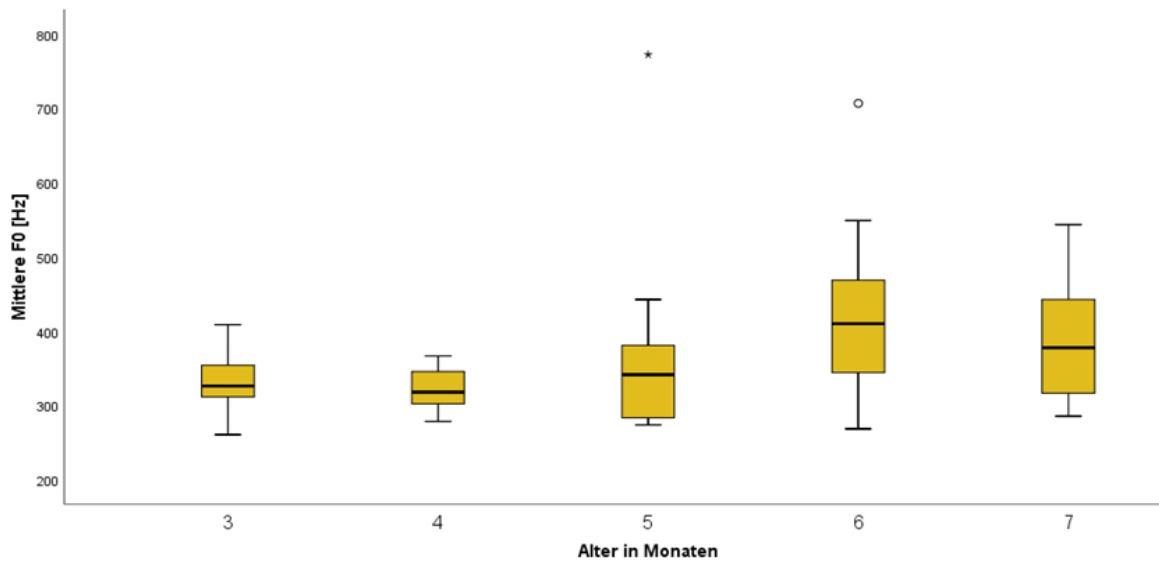


Abbildung 15: Boxplot für die Veränderung der mittleren F_0 [Hz] im Untersuchungszeitraum (kindbasiert)

Tabelle 12 liefert die entsprechenden Ergebnisse der kindbasierten Analyse für den absoluten F_0 -Range. Wie zu erwarten, ist der Boxplot in **Abbildung 16** ähnlich dem Bild in **Abbildung 15**. Der Friedmann-Test ergab signifikante Mittelwert-Unterschiede im Gesamtzeitraum ($p=0,01$). Ein entsprechender Post-hoc-Test (Dunn-Bonferroni Testung) konnte einen signifikanten Unterschied zwischen Monat 3 und Monat 6 aufdecken ($z= 2,83$ $p=0,02$ $r=0,90$). Der Korrelationskoeffizient (Pearson) von $r=0,90$ zeigt hier einen starken Effekt. Korrespondierend zur lautbasierten Analyse sind Spannweite und Interquartilabstand zum Zeitpunkt von Monat 5 und Monat 6 am größten.

Tabelle 12: Deskriptive Statistik des absoluten F_0 -Hubes pro Monat (kindbasiert)

Alter in Monaten		3	4	5	6	7	
		(n=11)	(n=11)	(n=10)	(n=10)	(n=6)	
Absoluter F_0 -Hub [Hz]	Mittelwert	128,6	137,3	184,9	238,1	176,1	
	SD	31,3	30,3	85,5	92	93,3	
	Median	131,9	135,4	154,6	230,8	156,5	
	Perzentile	25.	106,2	120,2	131,4	153,9	101,4
		75.	142,4	146,7	224,3	299,1	281,1
	Interquartilbereich	105,2	106,6	293,2	307	240,4	
	Minimum	93,3	89,5	91,8	124,6	55,1	
Maximum	198,5	196,61	385,6	431,6	295,5		

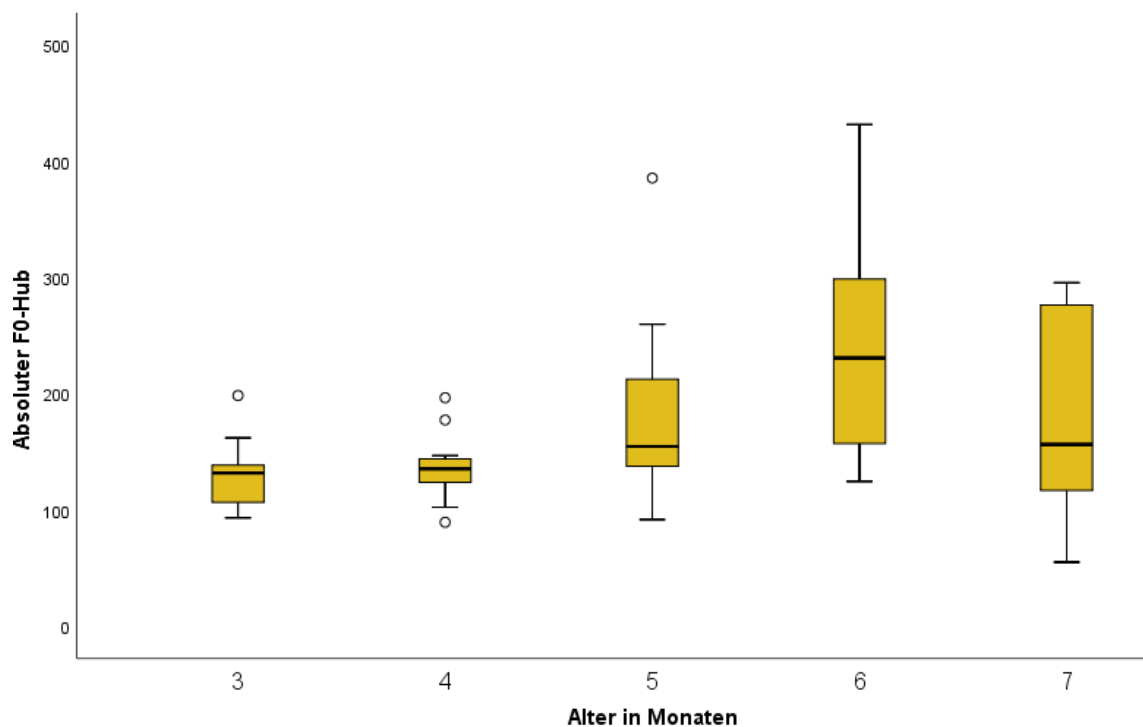


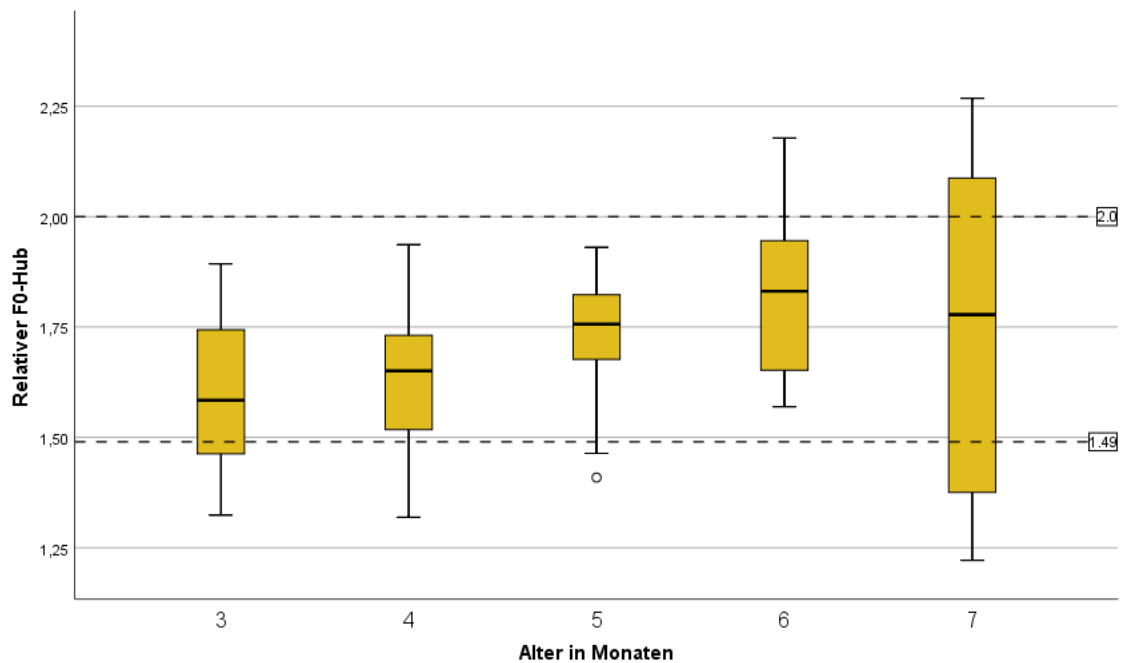
Abbildung 16: Boxplot für den absoluten F₀-Hub pro Monat (kindbasiert)

Tabelle 13 zeigt die Ergebnisse der kindbasierten Analyse für den relativen F₀-Hub. Der dazugehörige Boxplot in Abbildung 17 zeigt bezüglich der Verteilungseigenschaften hier einen deutlichen Anstieg im Untersuchungszeitraum. Die eingezeichneten Linien bei 1,49 und 2,0 verdeutlichen, dass die Kinder Komfortvokalisationen mit einem F₀-Range produzieren, der in einem Bereich zwischen einer musikalischen Quinte und einer Oktave liegen. Zum Zeitpunkt von 6 und 7 Monaten werden von einigen Kindern sogar Vokalisationen mit einem Frequenzverhältnis produziert, welches das einer musikalischen Oktave übersteigt.

Ähnlich wie bei der Analyse des absoluten F₀-Ranges ist ein kontinuierlicher Anstieg zu verzeichnen. Eine Varianzanalyse nach *Friedman* zeigte signifikante Mittelwertunterschiede ($p=0,03$) im Untersuchungszeitraum. Ein entsprechender Post-hoc-Test (Dunn-Bonferroni-Testung) konnte signifikante Unterschiede vom 3. Monat zum 6. Monat ($z= 2,67$ $p=0,04$ $r=0,84$) aufdecken. Der Korrelationskoeffizient (Pearson) von $r=0,84$ zeigt hier ebenfalls einen starken Effekt.

Tabelle 13: Deskriptive Statistik relativer F₀-Hub (kindbasierte Mittelung)

Alter in Monaten		3	4	5	6	7	
		(n=11)	(n=11)	(n=10)	(n=10)	(n=6)	
Relativer F ₀ -Hub	Mittelwert	1,60	1,63	1,72	1,83	1,75	
	SD	,21	,19	,17	,19	,40	
	Median	1,58	1,65	1,76	1,83	1,78	
	Perzentile	25.	1,43	1,49	1,62	1,64	1,34
		75.	1,88	1,75	1,84	1,97	2,13
	Interquartilbereich	,45	,26	,22	,33	,79	
	Minimum	1,32	1,32	1,41	1,57	1,22	
Maximum	1,89	1,94	1,93	2,18	2,27		

**Abbildung 17:** Boxplot für den relativen F₀-Hub (kindbasiert) mit horizontalen Orientierungslinien bei der musikalischen Quinte (1,49) und Oktave (2,0)

5.1.2 Analyse der Vokalisationslänge

5.1.2.1 Lautbasierte Analyse der Vokalisationslänge

Sämtliche Komfortvokalisationen, auch Kurzlaute, wurden bei der Charakterisierung der Vokalisationslänge berücksichtigt. **Tabelle 14** enthält Mittelwert, Standardabweichung, Median, Perzentile, Interquartil-Bereich sowie Minimum und Maximum der Vokalisationslängen im Untersuchungszeitraum.

Der Median der Vokalisationslänge liegt bei 634 ms. Die mittleren 50 Prozent der Vokalisationen (Interquartil-Bereich) weisen eine Länge zwischen 360 ms und 1084 ms auf (siehe Boxplot, **Abbildung 18**). Auch hier ergab die Untersuchung der statistischen Verteilung der Lautlängen keine hinreichende Normalverteilung (siehe Histogramm, **Abbildung 18**).

Tabelle 14: Deskriptive Statistik der Lautlänge

N = 5476		Statistik	
Vokalisationslänge [ms]	Mittelwert	833	
	Standardabweichung	670	
	Median	634	
	Perzentile	25.	360
		75.	1084
	Interquartil-Bereich	724	
	Minimum	45	
Maximum	6368		

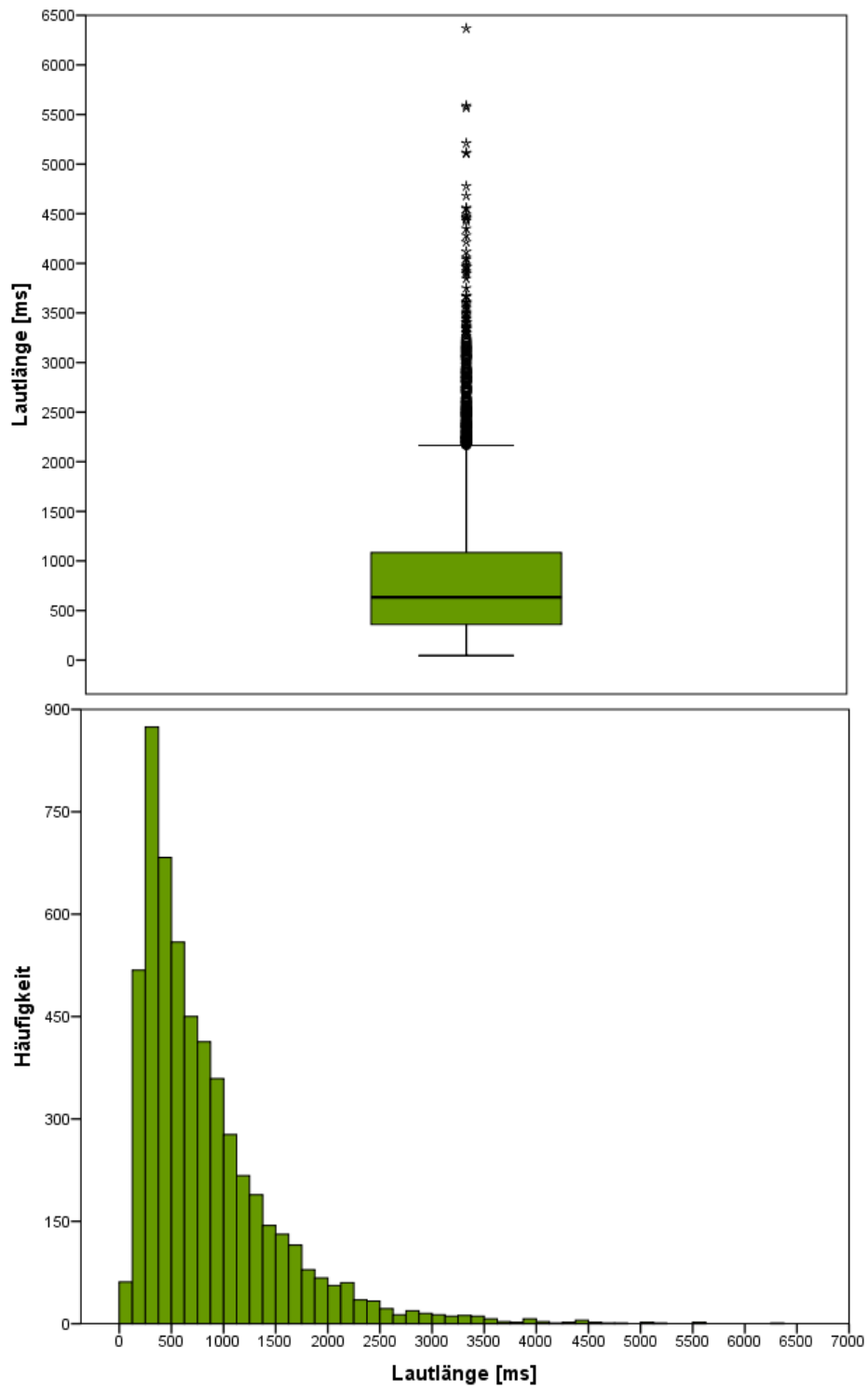


Abbildung 18: Boxplot (oben) und Histogramm (unten) der Vokalisationslänge [ms] im Gesamtzeitraum.

Tabelle 15 zeigt die statistischen Größen der Vokalisationslängen aller Vokalisierungen pro Monat. Es wird deutlich, dass die Vokalisationslänge bis zum Alter von 6 Monaten kontinuierlich zunimmt.

Der Median steigt zwischen dem 3. und 6. Monat von 440 ms auf 743 ms an (siehe Liniendiagramm, **Abbildung 20**). Auch Quartil-Abstand und Spannweite verbreitern sich im Entwicklungsverlauf (siehe Boxplot, **Abbildung 19**). Während die maximalen Lautlängen im Untersuchungszeitraum zunehmen, treten auch Laute mit weniger als 100 ms Lautlänge im gesamten Untersuchungszeitraum auf.

Tabelle 15: Deskriptive Statistik der Lautlänge [ms] pro Monat

Alter in Monaten		3 (n=863)	4 (n=1829)	5 (n=1237)	6 (n=1153)	7 (n=394)	
Vokalisationslänge [ms]	Mittelwert	615	779	877	994	949	
	SD	506	546	668	837	799	
	Median	440	633	715	743	707	
	Perzentile	25.	290	373	380	424	375
		75.	778	1030	1159	1305	1283
	Interquartil-Bereich	488	657	779	881	908	
	Minimum	65	45	45	95	71	
Maximum	3963	4428	5108	6368	4778		

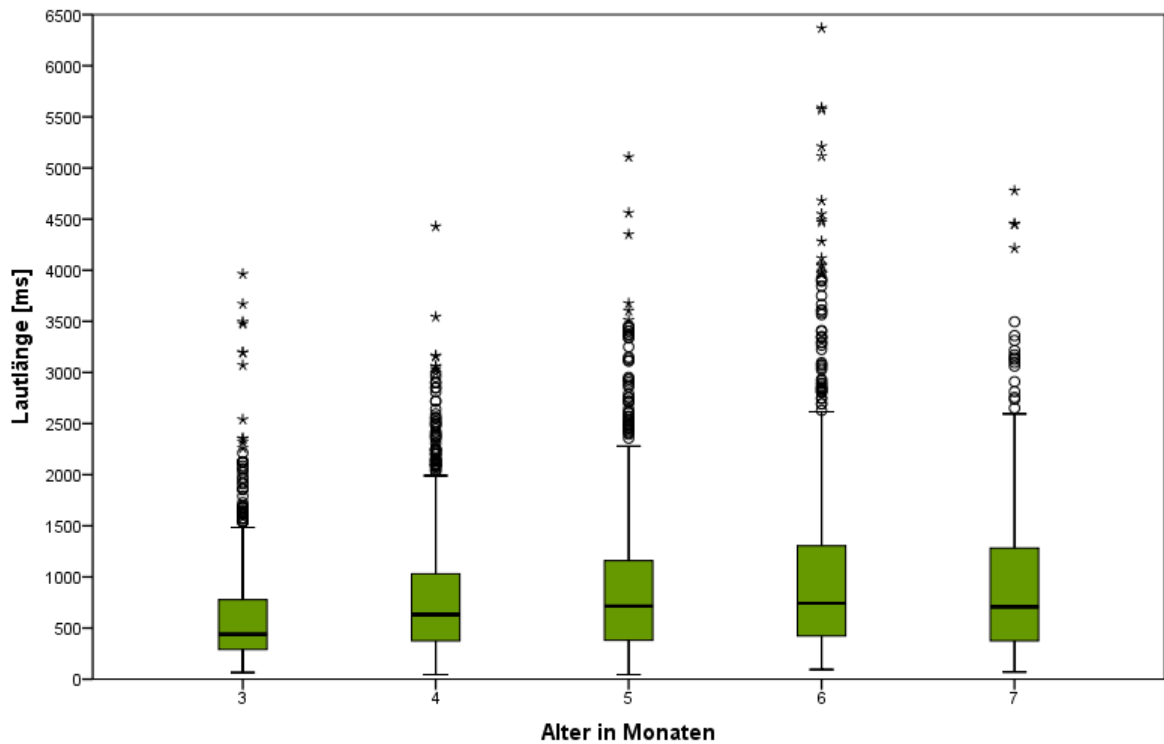


Abbildung 19: Boxplot für die Vokalisationslänge pro Monat

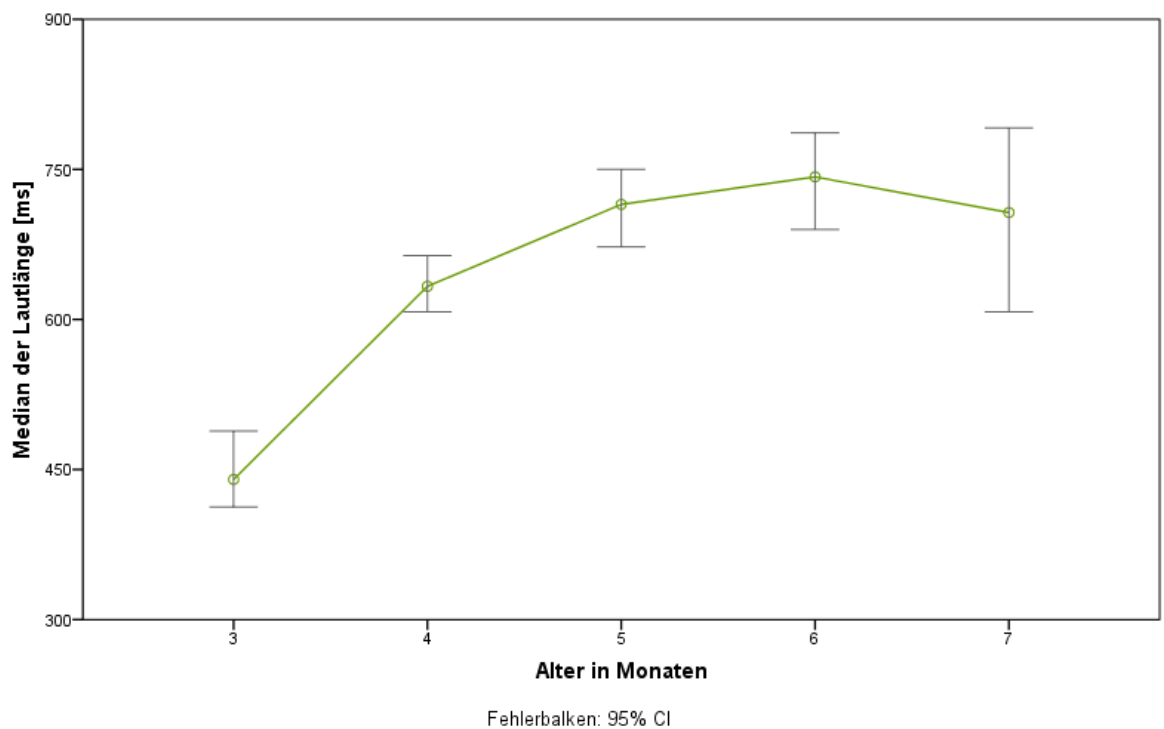


Abbildung 20: Liniendiagramm für den Median der Lautlänge [ms] pro Monat mit eingezeichneten 95%-Konfidenzintervallen (graue Whisker).

Der Boxplot in **Abbildung 21** zeigt die Verteilungseigenschaften der Vokalisationslänge der einzelnen Kinder im Gesamtzeitraum. Diese unterscheiden sich teilweise deutlich zwischen den Kindern.

Die Mediane der Kinder bewegen sich zwischen 481 ms und 747 ms (siehe **Tabelle 16**). Das Kind AD sticht mit einem Median von 1084 ms extrem aus der Gruppe heraus. Die Probandin weicht im Median mehr als zwei Drittel vom Gruppenmedian ab. Auch der Quartil-Abstand ist bei einigen Kindern ca. doppelt so breit wie bei anderen. Ähnliches gilt für die Maxima der Lautlängen. Die Minima der Lautlängen liegen bei den meisten Kindern (außer bei AC, VI, VE) unter 100 ms.

Tabelle 16: Deskriptive Statistik der mittleren Vokalisationslänge pro Kind

Lautlänge [ms] N= 5476		Mittelwert	SD	Median	Perzentile	Interquartil	Minimum	Maximum	
VB	♀	648	526	481	25.	295	538	45	3215
					75.	833			
VL	♀	664	501	521	25.	303	567	83	3611
					75.	870			
VF	♂	667	488	540	25.	330	520	45	3511
					75.	850			
AC	♂	737	448	592	25.	390	620	165	2996
					75.	1011			
VI	♂	757	557	630	25.	363	626	120	4470
					75.	990			
VD	♀	829	681	660	25.	325	769	65	5108
					75.	1095			
VG	♂	860	593	707	25.	461	663	71	6368
					75.	1124			
VH	♀	911	557	504	25.	285	1011	65	4778
					75.	1296			
AK	♂	956	747	695	25.	364	1034	90	3963
					75.	1398			
VE	♂	667	974	746	25.	450	963	105	5589
					75.	1413			
AD	♀	1175	693	1084	25.	657	921	76	3544
					75.	1578			

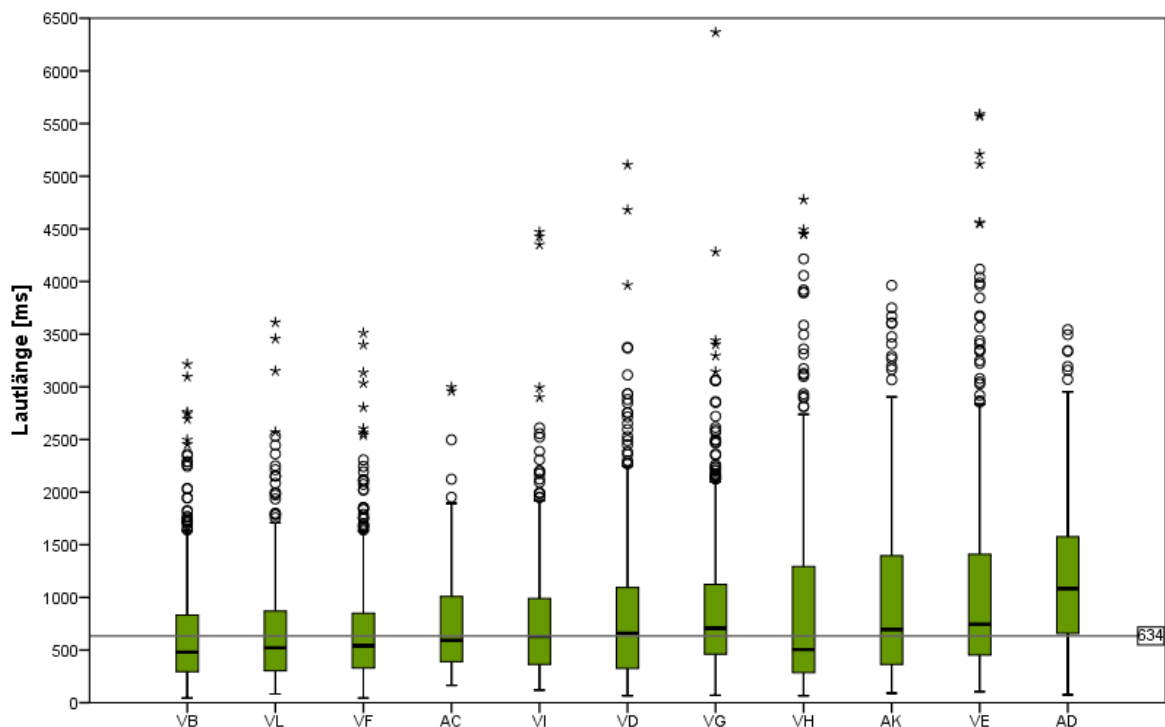


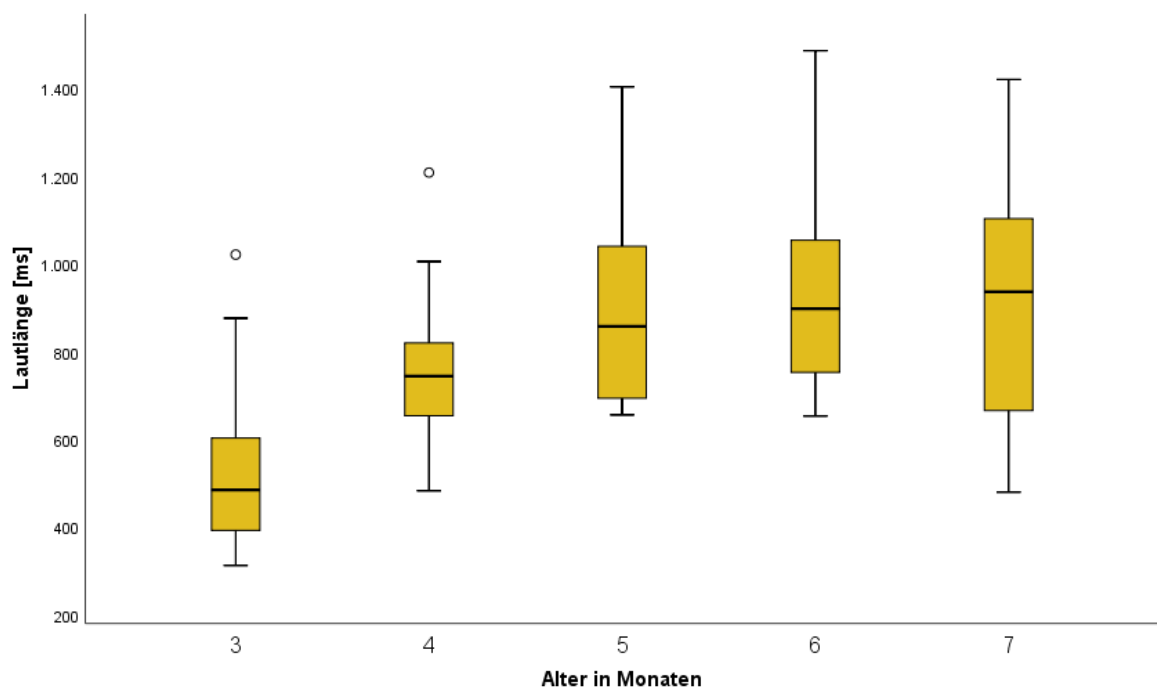
Abbildung 21: Boxplot der Vokalisationslänge [ms] pro Kind mit eingezeichnetem Gruppenmedian (graue Linie)

5.1.2.2 Kindbasierte Analyse der Vokalisationslänge

Tabelle 17 liefert die Ergebnisse der kindbasierten Analyse der Vokalisationslänge pro Monat. Der Boxplot in **Abbildung 22** veranschaulicht, dass sich die Ergebnisse der vokalisationsbasierten Analyse (**Abbildung 20**) hier widerspiegeln. Die Vokalisationslänge steigt kontinuierlich an. Eine zweifaktorielle Varianzanalyse nach Friedman zeigte signifikante Mittelwertunterschiede ($p=0,02$) im Untersuchungszeitraum. Ein entsprechender Post-hoc-Test (Dunn-Bonferroni-Testung) ergab signifikante Mittelwertunterschiede mit einer starken Effektstärke zwischen Monat 3 und Monat 6 ($z=2,67$ $p=0,04$ $r=0,84$). Zwischen Monat 3 und Monat 7 konnte ebenfalls ein starker Effekt ($r=0,79$) berechnet werden. Die Verteilung der Daten wird auch hier ab dem 5. Monat größer.

Tabelle 17: Deskriptive Statistik für die mittlere Vokalisationslänge pro Monat (kindbasierte Mittelung)

Alter in Monaten		3	4	5	6	7	
		(n=11)	(n=11)	(n=10)	(n=10)	(n=6)	
Vokalisationslänge [ms]	Mittelwert	546,1	771,3	898,6	971,7	923,6	
	SD	225,8	199,6	230,9	292,9	339,8	
	Median	485	744,2	857,7	898	936,4	
	Perzentile	25.	368,3	639	688,2	738,8	619,3
		75.	620,6	857,7	1041,9	1156,5	1182,3
	Interquartil-Bereich	252,3	218,7	353,7	417,7	563	
	Minimum	313,1	482,9	656,1	653,3	480	
	Maximum	1021,2	1208	1403,2	1485,7	1420,3	

**Abbildung 22:** Boxplot für die mittlere Vokalisationslänge pro Monat (kindbasiert)

5.2 Charakterisierung der Melodie-Struktur

Die nachfolgenden Tabellen zeigen die Häufigkeitsverteilung der einzelnen Melodie-Struktur-Typen im gesamten Untersuchungszeitraum (**Tabelle 18**) sowie eine Zusammenfassung einfacher, komplexer und segmentierter Melodiestructuren im zeitlichen Verlauf pro Monat (**Tabelle 19**). Betrachtet man das dazugehörige Balkendiagramm (**Abbildung 23**) so wird deutlich, dass die Produktion einfacher Strukturen (KL, 1B) im gesamten Untersuchungszeitraum dominiert. Vokalisationen mit doppelter Bogenstruktur treten mit 18,2% am zweithäufigsten auf. Einfache segmentierte Lautstrukturen (1S) machen etwa 10% aller Laute aus. Komplexere Strukturen (3B, MB, KS, 2S, 3S, MS) sind im Lautrepertoire der Probandengruppe nur sehr selten vertreten (jeweils $\leq 6\%$).

Allerdings nimmt der prozentuale Anteil einfacher Strukturen zwischen dem 3. und 7. Monat kontinuierlich zu Gunsten komplexerer Strukturen ab. Während die Produktion von einfachen Strukturen im Untersuchungszeitraum um etwas mehr als ein Viertel zurückgeht, steigt die Produktion von Mehrfachbögen und segmentierten Lauten um fast das Doppelte an.

Tabelle 18: Häufigkeiten einzelner Melodie-Struktur-Typen im Untersuchungszeitraum

Lautstruktur	Absolute Häufigkeit	Relative Häufigkeit [%]
KL	747	13,6
1B	2270	41,5
2B	998	18,2
3B	328	6,0
MB	174	3,2
1S	562	10,3
2S	122	2,2
3S	24	0,4
KS	247	4,5
MS	4	0,1
Gesamt	5476	100,0

Tabelle 19: Auftrittshäufigkeiten für einfache, komplexe und segmentierte Lautstrukturen pro Monat

Alter in Monaten	Lautstruktur	Häufigkeit	Prozent
3	1B	588	68,1
	MB	154	17,8
	S	121	14,0
	Gesamt	863	100,0
4	1B	1016	55,5
	MB	503	27,5
	S	310	16,9
	Gesamt	1829	100,0
5	1B	644	52,1
	MB	368	29,7
	S	225	18,2
	Gesamt	1237	100,0
6	1B	581	50,4
	MB	353	30,6
	S	219	19,0
	Gesamt	1153	100,0
7	1B	188	47,7
	MB	122	31,0
	S	84	21,3
	Gesamt	394	100,0

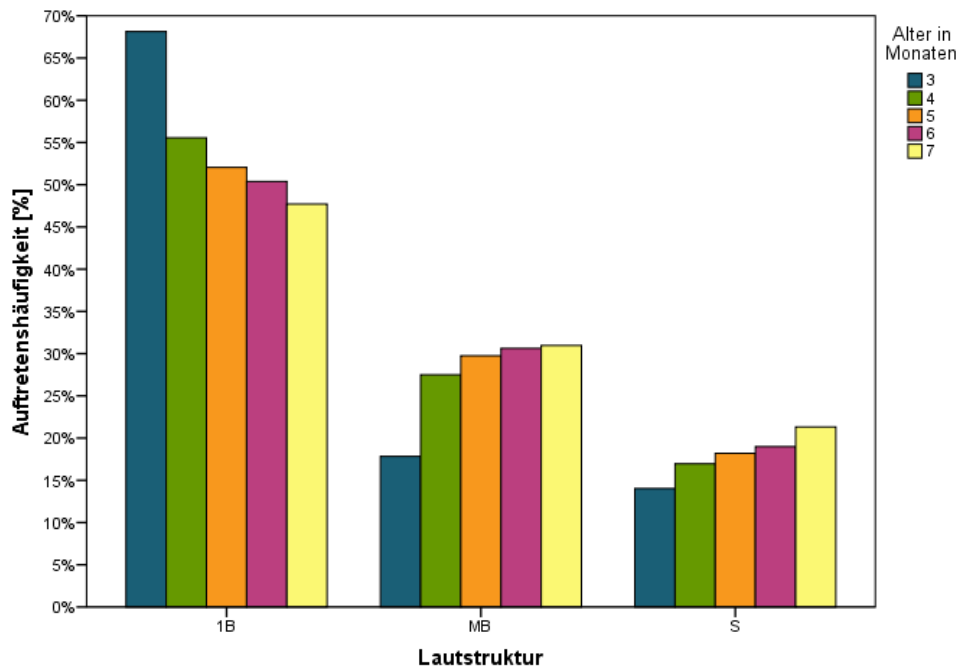


Abbildung 23: Verteilung der zusammengefassten Melodiestructuren (Einfachbögen+KL, Mehrfachbögen, segmentierte Laute) pro Monat

5.3 Individuelle Variation der Melodiestructur - der Melodiekomplexitätsindex (MCI)

Wie in Kapitel 3 und 4 beschrieben, wurde neben einer Häufigkeitsanalyse der einzelnen Vokalisationsstrukturen auch der Melodie-Komplexitäts-Index (MCI) als Maßzahl für die kindbasierte Analyse ermittelt. Er stellt den Anteil der Laute mit komplexen Melodiestructuren an der Gesamtzahl der Vokalisationen dar. Dies erfolgte auf Basis der einzelnen Vokalisationen der Kinder pro Aufnahmetag. **Tabelle 20** zeigt zunächst die deskriptive Statistik des MCI's gemittelt über alle Aufnahmetage im Untersuchungszeitraum. Der Mittelwert von 0,45 bedeutet, dass die komplexen Melodiestructuren im Durchschnitt einen Anteil von 45% an der Gesamtanzahl der Vokalisationen haben. Der MCI variiert von 0 (Min) bis 1 (Max) und ist somit sehr groß. Der Boxplot in **Abbildung 25** verdeutlicht, dass die Mediane der ersten sechs Kinder unter dem Gruppenmedian von 0,42 liegen, während die restlichen Kinder höhere Medianwerte aufweisen (siehe auch **Tabelle 21**). Die Kinder VF und AC sind die einzigen im Gesamtzeitraum, deren Medianwerte unter dem Gruppenmedian liegen und gleichzeitig keinen Maximalwert von über 0,7 erreichen. Das Kind AD hingegen sticht mit seinem überdurchschnittlich hohen Median von 0,79 und seinem Maximum von 1,00 deutlich aus der Gruppe heraus. Bei dem Minimum von 0,08 soll erwähnt werden, dass es sich dabei um einen von einer einzigen Aufnahme stammenden MCI-Wert vom Kind VD handelt und die Interpretation dieses Wertes möglicherweise mit Unsicherheit behaftet ist.

Tabelle 20: Deskriptive Statistik für den MCI (alle Aufnahmetage gehen ein)

N = 131 (pro Aufnahmetag)		Statistik	
Melody Complexity Index	Mittelwert	,45	
	Standardabweichung	,23	
	Median	,42	
	Perzentile	25.	,29
		75.	,62
	Interquartil-Bereich	,33	
	Minimum	,00	
	Maximum	1,00	

Tabelle 21: Deskriptive Statistik für den MCI pro Kind (alle Aufnahme tage gehen ein)

MCI	Mittelwert	SD	Median	Perzentile		Interquartil	Minimum	Maximum
				25.	75.			
VF ♂ n=12	,36	,13	,35	,24	,44	,20	,14	,55
VI ♂ n=12	,38	,18	,38	,23	,55	,32	,11	,71
AC ♂ n=6	,39	,13	,35	,31	,50	,19	,23	,61
VH ♀ n=22	,40	,24	,40	,20	,62	,42	,00	,77
VL ♀ n=14	,42	,20	,40	,23	,62	,39	,12	,75
VB ♀ n=12	,42	,19	,37	,29	,54	,25	,12	,79
VG ♂ n=14	,43	,17	,48	,29	,53	,24	,16	,69
AK ♂ n=9	,46	,21	,53	,26	,65	,39	,18	,71
VE ♂ n=8	,47	,17	,52	,30	,62	,32	,22	,64
VD ♀ n=13	,56	,26	,57	,35	,77	,42	,08	1,00
AD ♀ n=9	,80	,15	,79	,64	,94	,30	,59	1,00

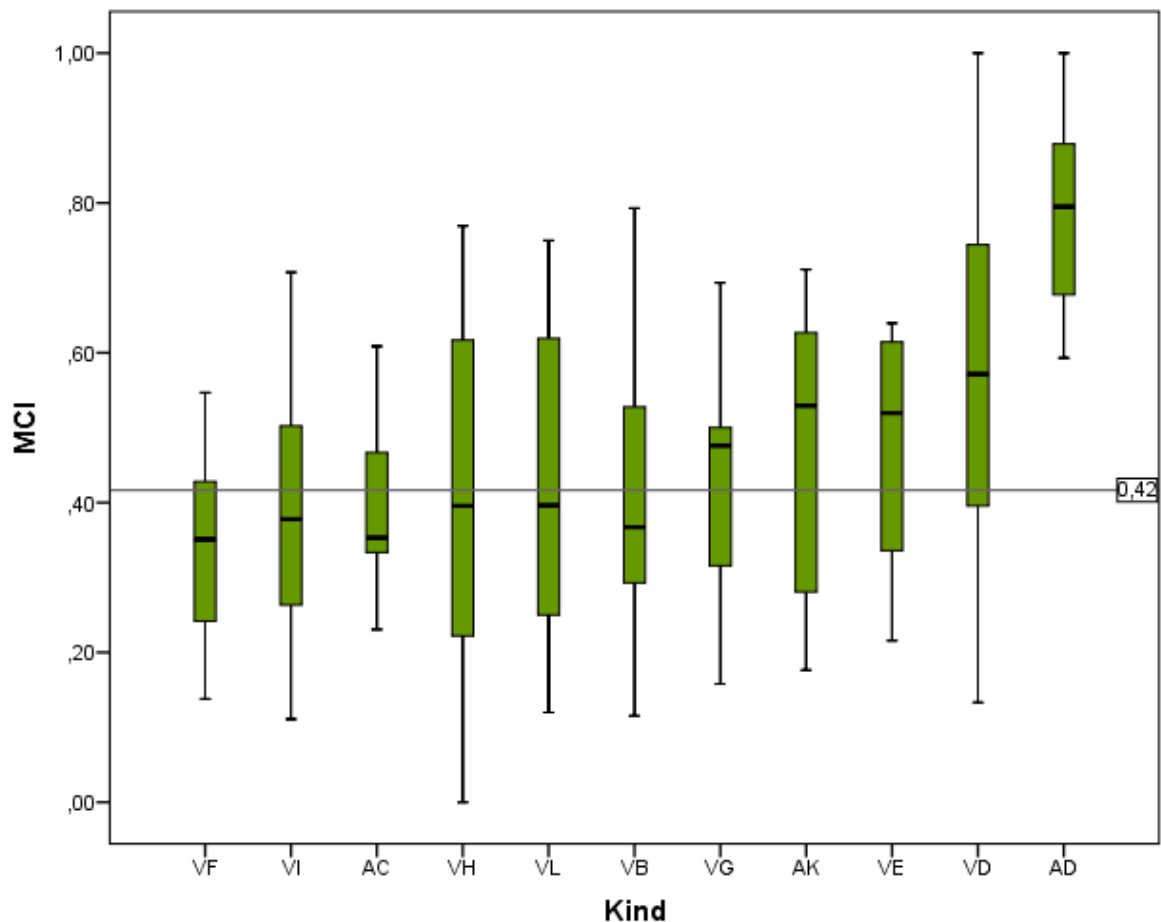


Abbildung 24: Verteilung der MCI-Werte pro Kind (es gehen alle Aufnahmetage pro Kind ein) mit eingezeichnetem Gruppenmedian (graue Linie)

Bei der Analyse der Melodiekomplexität im Entwicklungsverlauf geht pro Monat jeweils ein Durchschnittswert pro Kind ein (siehe **Tabelle 22**). Mit Ausnahme des 6. Monats steigen die Mediane im Untersuchungszeitraum kontinuierlich an. Zwischen dem 3. und 4. Monat scheint es den größten Entwicklungssprung zu geben. Hier steigt der Median fast auf das Doppelte an. Die Verteilung der Werte ist zwischen den einzelnen Monaten sehr unterschiedlich, allerdings steigen auch die Minimum-Werte kontinuierlich an, sodass davon auszugehen ist, dass die Produktion von einfachen Melodiestructuren nicht nur bei einzelnen Kindern, sondern in der Gesamtgruppe abnimmt.

Der Boxplot in **Abbildung 24** zeigt, dass das Kind VD insgesamt betrachtet, weit über dem Gruppendurchschnitt liegt. Der untere Ausreißer im 4. Monat stammt von dem Kind VH. Auch dieses Kind zeigt ansonsten eine sehr breite Verteilung der MCI-Werte (siehe Boxplot in **Abbildung 24**).

Die Minima im 5., 6., und 7. Monat bewegen sich zwischen 0,23 und 0,43 und stammen ebenfalls alle von unterschiedlichen Probanden. Die Maxima vom 3.-6. Monat stammen

hingegen alle von dem Kind VD. Das Maximum im 7. Monat ist auf das Kind VB zurückzuführen (von VD liegen nur Aufnahmen bis einschließlich zum 6. Monat vor).

Tabelle 22: MCI pro Monat (ein Monatswert pro Kind)

Alter in Monaten		3 (n=11)	4 (n=11)	5 (n=10)	6 (n=10)	7 (n=5)	
Melody Complexity Index	Mittelwert	,27	,44	,49	,46	,60	
	SD	,18	,16	,20	,11	,12	
	Median	,22	,42	,51	,42	,61	
	Perzentile	25.	,17	,36	,35	,40	,48
		75.	,31	,47	,55	,51	,67
	Interquartil-Bereich	,14	,11	,20	,10	,19	
	Minimum	,08	,12	,23	,36	,43	
	Maximum	,72	,77	,94	,72	,79	

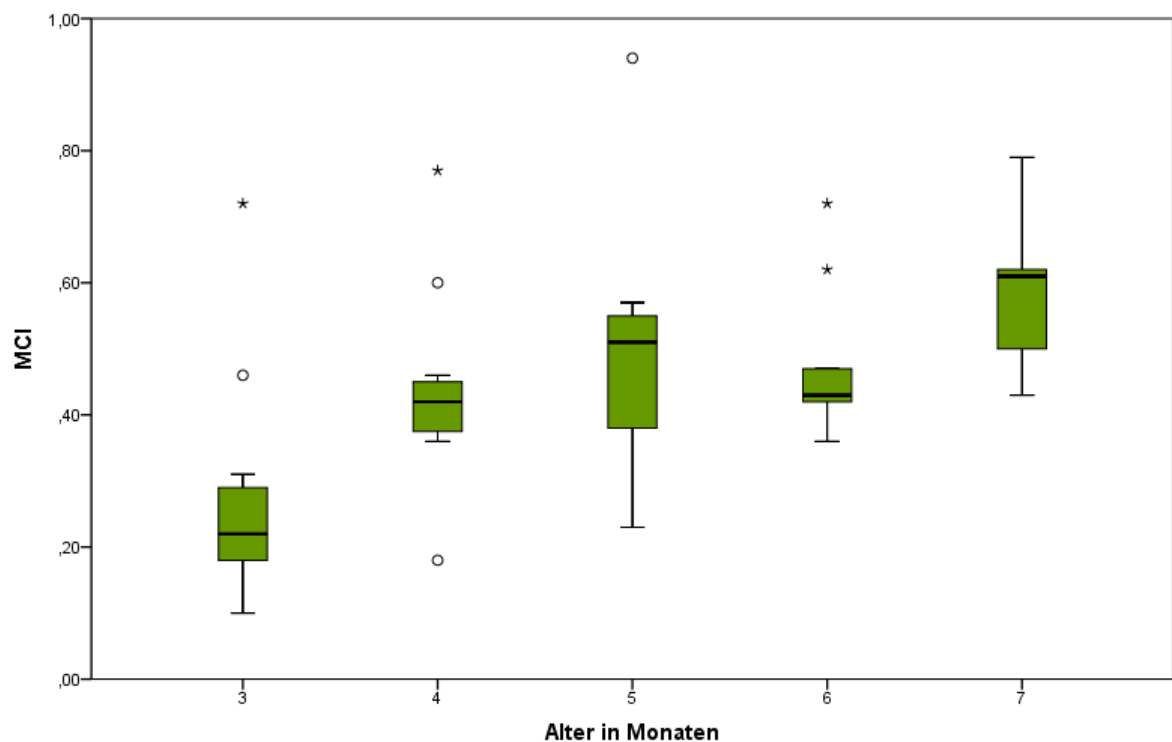


Abbildung 25: Boxplot für den MCI pro Monat (pro Monat geht von jedem Kind ein Wert ein)

Die statistische Analyse lieferte durch den Friedman-Test signifikante Unterschiede für den MCI im Untersuchungszeitraum ($p=0,01$). Ein entsprechender Post-hoc-Test (Dunn-Bonferroni-Testung) zeigte signifikante Unterschiede zwischen Monat 3 und Monat 5 ($z=2,90$; $p=0,04$; Effektstärke nach Pearson: $r=0,92$) sowie zwischen Monat 3 und Monat 7 ($z=3,10$; $p=0,02$; Effektstärke nach Pearson: $r=0,98$). Starke, jedoch nicht signifikante Effekte zeigen sich auch zwischen Monat 3 und Monat 6 ($z=2,2$; $p=0,28$; $r=0,70$) sowie zwischen Monat 3 und Monat 4 ($z=1,80$; $p=0,72$; $r=0,57$). Es konnten keine signifikanten Geschlechtsunterschiede gefunden werden.

5.4 Charakterisierung der artikulatorischen Fähigkeiten

Neben einer eher larynx-fokussierten Analyse der stimmphysiologischen Parameter und der Melodiekomplexität, wurden parallel artikulatorische Aktivitäten auf der supralaryngealen Ebene analysiert. Um mögliche Beziehungen zwischen der melodischen und artikulatorischen Entwicklung aufzudecken, erfolgte die Analyse von supralaryngealen Aktivitäten nicht nur auf Vokalisationsebene, sondern auch auf Ebene einzelner Melodiestrukturen.

5.4.1 Die relative Häufigkeit artikulierter Vokalisationen

Wie aus Kapitel 4 hervorgeht, gingen insgesamt 5476 Komfortvokalisationen in diese spezifische Analyse ein. Von diesen 5476 Vokalisationen sind 2130 Vokalisationen artikulierte, d.h. sie weisen mindestens ein artikulatorisches Element auf. Dies entspricht 39% aller Vokalisationen im Untersuchungszeitraum.

Dieses Verhältnis von artikulierten zu nicht artikulierten Vokalisationen verändert sich im Untersuchungszeitraum vor allem vom 3. zum 4. Monat (siehe **Tabelle 23**). Es ist ein leichter Zuwachs des Anteils artikulierter Vokalisationen zu beobachten, welcher sich mit einem Peak im 6. Monat zwischen 6% und 13% bewegt. Es wurden keine signifikanten Geschlechtsunterschiede im Verhältnis von artikulierten zu nicht artikulierten Vokalisationen gefunden.

Tabelle 23: Häufigkeiten der Vokalisationen mit und ohne Artikulation pro Monat und Gesamtzeitraum

			Alter in Monaten					Gesamt
			3	4	5	6	7	
Artikulation	ja	Anzahl	256	756	481	497	140	2130
		Häufigkeit [%]	29,7%	41,3%	38,9%	43,1%	35,5%	38,9%
	nein	Anzahl	607	1073	756	656	254	3346
		Häufigkeit [%]	70,3%	58,7%	61,1%	56,9%	64,5%	61,1%

5.4.2 Die Komplexität der Artikulation - Anzahl artikulatorischer Elemente pro Vokalisation

Um korrespondierend zur Melodiekomplexität (Kapitel 5.3) zu untersuchen, wie komplex die Artikulation bezogen auf die Einzelvokalisation im Untersuchungszeitraum ist, wurde ermittelt wie viele artikulatorische Elemente im Durchschnitt pro Vokalisation produziert werden (siehe **Tabelle 24**).

Die artikulierten Vokalisationen (N=2130) enthielten zu ca. 70% nur ein artikulatorisches Element. 20% der artikulierten Vokalisationen enthielten zwei artikulatorische Elemente. Vokalisationen, in denen drei und mehr artikulatorische Elemente produziert werden traten nur mit einer Häufigkeit von 10% auf (siehe Kreisdiagramm **Abbildung 26**).

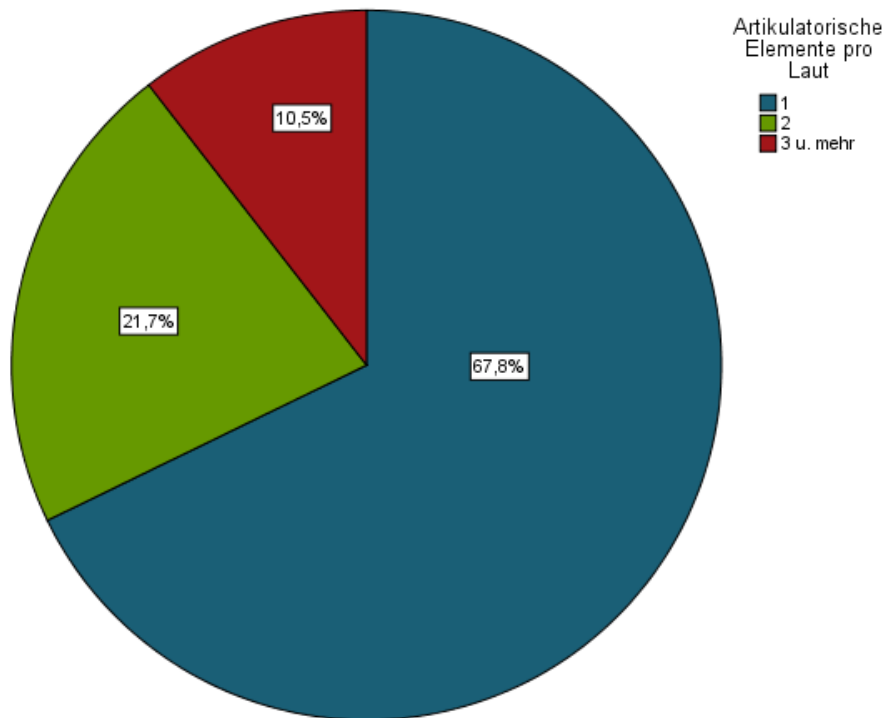


Abbildung 26: Kreisdiagramm für die Häufigkeitsverteilung der Laute mit einem, zwei sowie drei und mehr artikulatorischen Elementen im Gesamtzeitraum

Betrachtet man die Komplexität der Artikulation pro Monat, so lässt sich eine deutliche Entwicklungstendenz erkennen (siehe Balkendiagramm, **Abbildung 27**). Während die Produktion von Vokalisationen mit einfacher Artikulation (1 artikulatorisches Element pro Vokalisation) vom 3. bis zum 7. Monat kontinuierlich abnimmt, nimmt die Produktion von Vokalisationen mit komplexer Artikulation (2, 3 und mehr artikulatorische Elemente pro Vokalisation) im Entwicklungsverlauf stark zu. Laute mit 3 und mehr artikulatorischen Elementen werden im 7. Monat sogar fünfmal öfter produziert als noch im 3. Monat.

Tabelle 24: Häufigkeitsverteilung der Vokalisationen mit einem, zwei sowie drei und mehr artikulatorischen Elementen pro Monat und Gesamtzeitraum

Anzahl artikulatorischer Elemente pro Vokalisation	Alter in Monaten	Häufigkeit	Prozent
1	3	210	82,0
	4	542	71,7
	5	321	66,7
	6	295	59,4
	7	76	54,3
	Gesamt	1444	67,8
2	3	36	14,1
	4	147	19,4
	5	116	24,1
	6	129	26,0
	7	34	24,3
	Gesamt	462	21,7
3 u mehr	3	10	3,9
	4	67	8,9
	5	44	9,1
	6	73	14,7
	7	30	21,4
	Gesamt	224	10,5

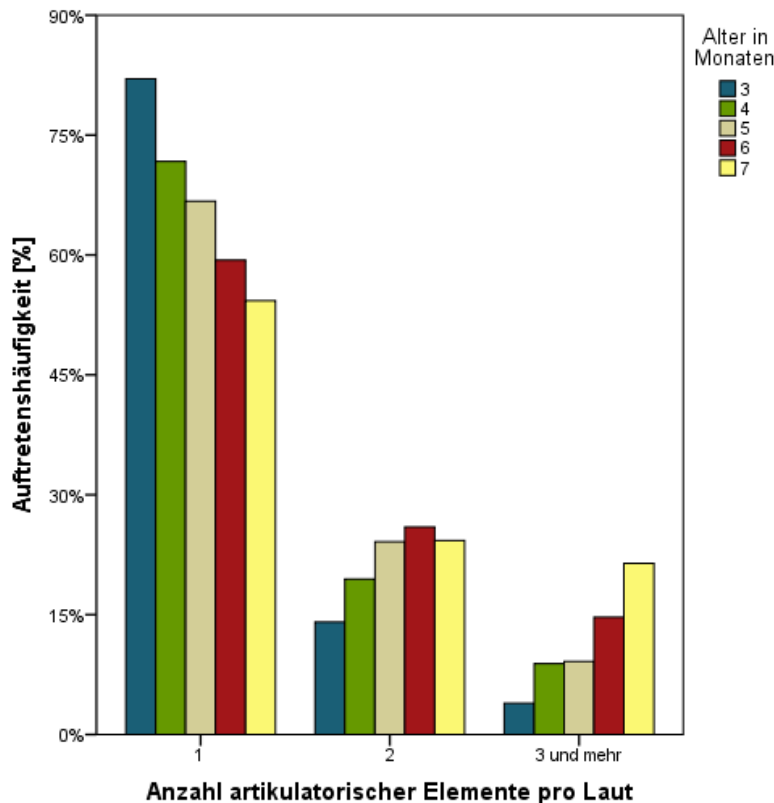


Abbildung 27: Balkendiagramm für die Häufigkeitsverteilung der Vokalisationen mit einem, zwei sowie drei und mehr artikulatorischen Elementen pro Monat

5.4.3 Die Produktion artikulatorischer Elemente

5.4.3.1 Vokalisationsbasierte Analyse artikulatorischer Elemente

Das Kreisdiagramm (**Abbildung 28**) mit zugehöriger **Tabelle 25** veranschaulicht die Gesamtverteilung der produzierten artikulatorischen Elemente im Untersuchungszeitraum. Insgesamt wurden im Untersuchungszeitraum 3186 (100%) artikulatorische Elemente produziert. Ein Drittel davon wurde dorsal, also im hinteren Bereich des Vokaltraktes gebildet (30,7%). Fast genauso oft wurden Labiale (25,6%) produziert. Klar wahrnehmbare Resonanzveränderungen (z.B. eine Veränderung von einem geschlossenen zu einem offenen Vokal) traten mit einer Häufigkeit von 20,9% auf. Apikale (im zentralen Vokaltrakt gebildet) traten mit 13,1% im Untersuchungszeitraum noch relativ selten auf, während laryngeal gebildete Artikulationen mit 3,6% kaum noch produziert wurden. In 6,2% der Fälle konnte der Artikulationsort innerhalb des Vokaltraktes nicht genau zugeordnet werden (Fragezeichen, unklar).

Tabelle 25: Häufigkeitsverteilung unterschiedlicher artikulatorischer Elemente im Untersuchungszeitraum

Artikulation	Häufigkeit	Prozente
labial	810	25,4
apikal	417	13,1
rv	667	20,9
dorsal	979	30,7
laryngeal	116	3,6
unklar	197	6,2
Gesamt	3186	100,0

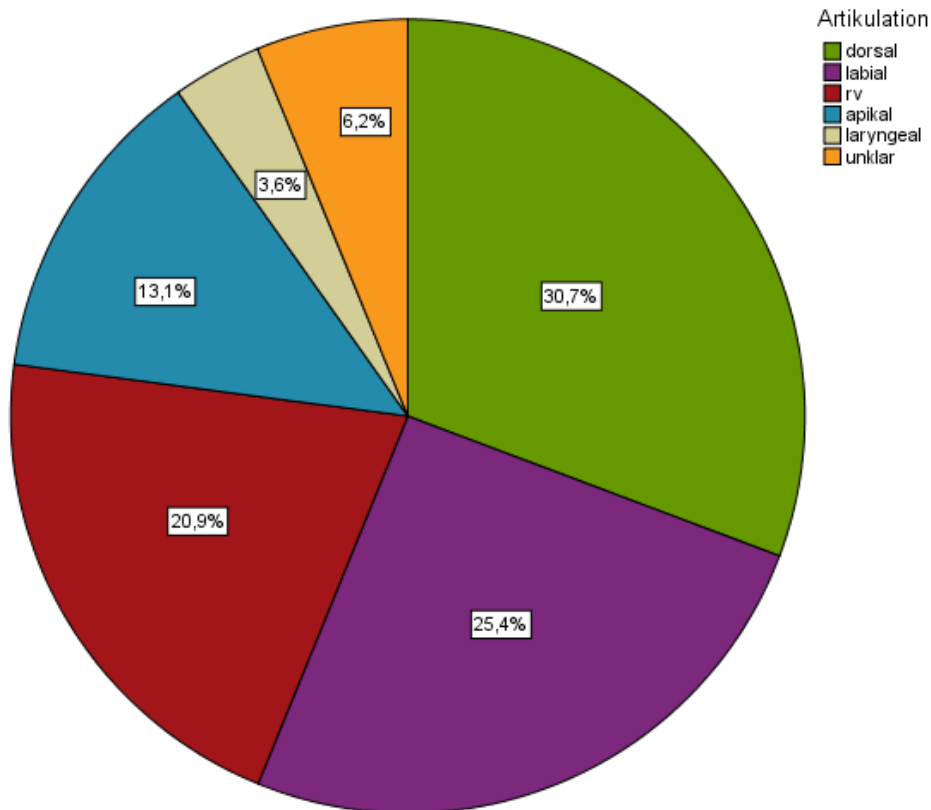


Abbildung 28: Kreisdiagramm für die Häufigkeitsverteilung einzelner artikulatorischer Elemente im Untersuchungszeitraum.

Das Balkendiagramm in **Abbildung 29** und dazugehörige **Tabelle 26** zeigen einen deutlichen Entwicklungstrend hinsichtlich der Verteilung der verschiedenen artikulatorischen Elemente im Laufe des Untersuchungszeitraumes. Der Anteil dorsal

gebildeter artikulatorischer Elemente geht zwischen dem 3. und 7. Monat um ca. drei Viertel zurück²³. Der Anteil labial produzierter artikulatorischer Elemente hingegen steigt um zwei Drittel an²⁴. Die Häufigkeit der apikal produzierten Elemente verdoppelt sich im Untersuchungszeitraum sogar. Die Produktion von reinen Resonanzveränderungen ohne konsonanten-ähnliche Verschlussbildung nimmt im Gegenzug um etwa die Hälfte ab. Die Produktion laryngeal gebildeter Elemente verändert sich im Untersuchungszeitraum nur unwesentlich. Auch der Anteil an Vokalisationen, bei denen der Artikulationsort nicht genau zugeordnet werden konnte verhält sich zwischen dem 3. und 7. Monat stabil.

Tabelle 26: Anzahl und Häufigkeiten einzelner artikulatorischer Elemente pro Monat

Artikulation		Alter in Monaten					Gesamt
		3	4	5	6	7	
dorsal	Anzahl	126	465	199	166	23	979
	%	39,9%	43,4%	27,9%	20,1%	8,9%	30,7%
labial	Anzahl	52	226	144	267	121	810
	%	16,5%	21,1%	20,2%	32,3%	46,7%	25,4%
apikal	Anzahl	30	75	106	148	58	417
	%	9,5%	7,0%	14,8%	17,9%	22,4%	13,1%
rv	Anzahl	73	212	194	157	31	667
	%	23,1%	19,8%	27,2%	19,0%	12,0%	20,9%
laryngeal	Anzahl	10	35	35	28	8	116
	%	3,2%	3,3%	4,9%	3,4%	3,1%	3,6%
unklar	Anzahl	25	58	36	60	18	197
	%	7,9%	5,4%	5,0%	7,3%	6,9%	6,2%
Gesamt	Anzahl	316	1071	714	826	259	3186
	%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%

²³ Der leichte Anstieg zwischen Monat 3 und Monat 4 ist möglicherweise auf die vergleichsweise geringe Fallzahl im dritten Monat zurückzuführen.

²⁴ Der leichte Abfall zwischen Monat 4 und Monat 5 ist möglicherweise auf den großen Unterschied in der Anzahl der Fälle zwischen beiden Monaten zurückzuführen.

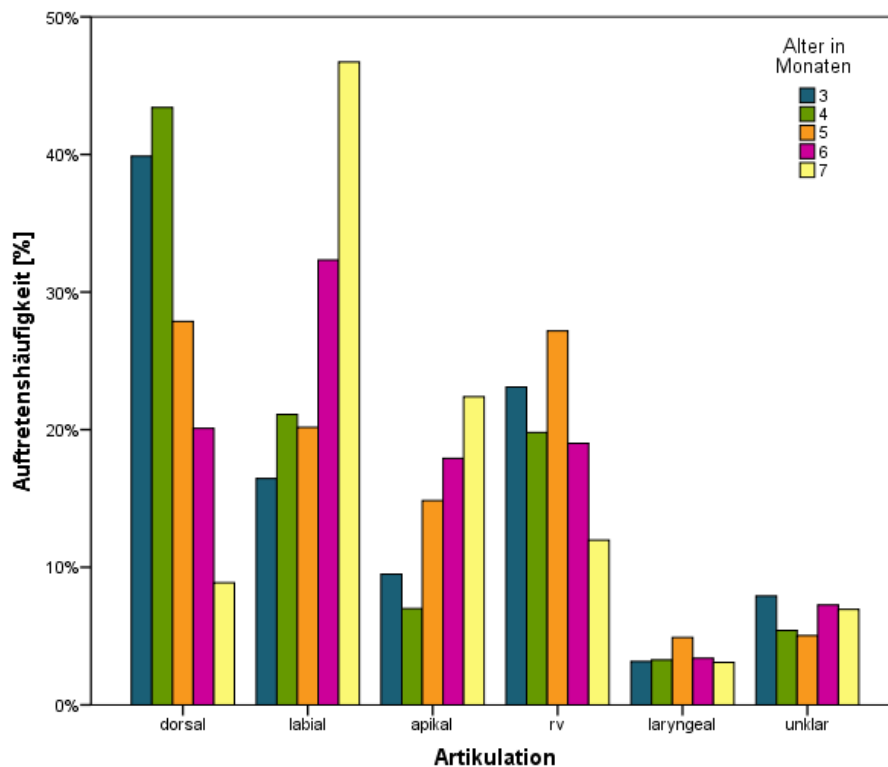


Abbildung 29: Balkendiagramm für die Häufigkeiten einzelner artikulatorischer Elemente pro Monat

5.4.3.2 Kindbasierte Analyse artikulatorischer Elemente

Der Boxplot in **Abbildung 30** stellt die Häufigkeitsverteilung der einzelnen artikulatorischen Elemente kindgemittelt dar. Für jedes Kind fließt ein Gesamtwert pro Artikulationsort ein (**Tabelle 27**).

Tabelle 28 verdeutlicht, dass sich die Kinder in der Produktion von artikulatorischen Elementen teilweise stark unterscheiden. Am größten ist die Spannweite bei der dorsalen Artikulation. Es treten sowohl Häufigkeiten von über 50% als auch von unter 10% bei den einzelnen Kindern auf.

Ähnliches gilt für die Produktion von labialen konsonantischen Elementen. Während das artikulatorische Repertoire bei einigen Kindern teilweise weniger als 10% Labiale aufweist (AC, VF, VH), besteht es hingegen bei einem Kind bereits zu über 50% aus Labialen (AK). Bei den restlichen Kindern treten Labiale mit einer Häufigkeit zwischen etwa 20 und 30% auf.

Etwa gleich häufig treten Resonanzveränderungen bei den einzelnen Kindern auf, allerdings ist die Verteilung gleichmäßiger. Bei allen Kindern besteht das Artikulations-Repertoire zu mindestens 12% aus Resonanzveränderungen.

Die apikal gebildeten artikulatorischen Elemente werden wie weiter oben bereits ersichtlich insgesamt deutlich seltener produziert. Doch auch hier produzieren einige Kinder apikale Elemente um ein Vielfaches mehr als andere Kinder.

Die Häufigkeitsverteilung der Produktion von Laryngealen ist relativ eng und bewegt sich in etwa zwischen 1-6% für die einzelnen Kinder. Ähnlich verhält es sich mit der Verteilung der unklaren Artikulationsorte. Der Ausreißer mit einer Häufigkeit von fast 10% stammt von dem Kind VB und ist möglicherweise auf eine geringere Aufnahmequalität einzelner Aufnahmetage zurückzuführen.

Tabelle 27: Häufigkeitsverteilung einzelner artikulatorischer Elemente (kindgemittelt) im Untersuchungszeitraum (pro Kind geht ein Gesamtprozentwert pro Artikulationsort ein)

Aufretenshäufigkeit Artikulationsort [%]		dorsal (n=11)	labial (n=11)	apikal (n=11)	rv (n=11)	laryngeal (n=11)	unklar (n=11)
Mittelwert		28,9	23,9	13,3	23,8	4,0	6,0
SD		15,2	14,3	6,1	7,8	1,6	1,8
Median		30,3	22,5	12,1	22,2	4,3	5,6
Perzentile	25.	18,2	9,9	8,7	19,0	2,8	5,1
	75.	42,0	31,3	17,6	28,1	4,9	7,1
Interquartilbereich		23,8	21,4	8,9	9,1	2,1	2,0
Minimum		2,8	4,3	6,0	12,6	1,2	2,9
Maximum		53,2	56,4	26,2	38,4	6,7	9,6

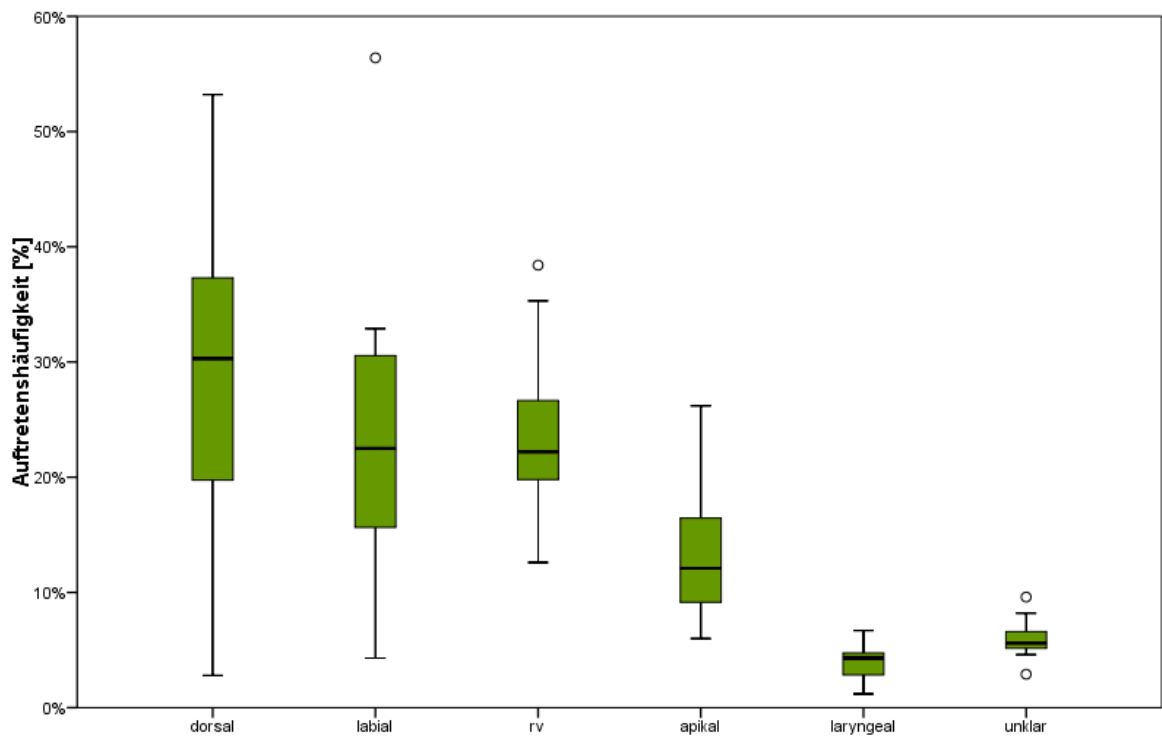


Abbildung 30: Boxplot für die Häufigkeitsverteilung einzelner artikulatorischer Elemente (kindgemittelt) im Untersuchungszeitraum (Gesamtprozentwerte der einzelnen Kinder gehen ein)

Tabelle 28: Relative Häufigkeit einzelner artikulatorischer Elemente pro Kind im Untersuchungszeitraum [%]

Kind	dorsal	labial	apikal	rv	laryngeal	unklar
AC	30,3	9,1	12,1	38,4	4,0	6,1
AD	42,0	21,4	6,0	20,6	4,5	5,6
AK	2,8	56,4	8,7	25,2	2,3	4,6
VB	21,3	31,3	9,6	22,2	6,1	9,6
VD	18,2	29,8	20,1	23,9	2,9	5,1
VE	24,0	22,5	26,2	15,4	4,9	7,1
VF	48,8	9,9	10,6	21,3	1,2	8,2
VG	32,6	32,9	13,0	12,6	2,8	6,1
VH	53,2	4,3	7,2	28,1	4,3	2,9
VI	13,1	24,2	17,6	35,3	4,6	5,2
VL	31,9	21,5	15,3	19,0	6,7	5,5

5.4.4 Produktion artikulatorischer Elemente in Abhängigkeit von der Melodiestructur

Betrachtet man die Produktion artikulatorischer Elemente in Abhängigkeit von der Melodiestructur, so wird deutlich, dass zunächst hauptsächlich Vokalisationen artikuliert werden, welche eine einfache Melodiestructur aufweisen (**Tabelle 29**).

Mit zunehmendem Alter nimmt die Produktion von artikulatorischen Elementen in Vokalisationen mit einfacher Melodiestructur kontinuierlich ab, während die Produktion von artikulatorischen Elementen in Vokalisationen mit komplexen Melodiestructuren (MB, S) kontinuierlich zunimmt (**Abbildung 31, Tabelle 29**). Im 3. Lebensmonat weisen fast 60% der artikulierten Laute eine einfache Melodiestructur auf. Vokalisationen mit mehrböigen oder segmentierten Melodiestructuren machen somit nur einen Anteil von etwa 40% aller artikulierten Vokalisationen aus. Zwischen dem 3. und 4. Monat kommt es zu einem deutlichen Anstieg von artikulierten Lauten mit mehrböigen Strukturen um 13%, während die artikulierten einfachen Melodiestructuren um fast 20 % abnehmen. Zwischen dem 6. und 7. Monat kommt es erneut zu einem deutlichen Abfall von 13% im Auftreten von artikulierten einfachen Melodiestructuren. Parallel dazu vollzieht sich zum gleichen Zeitpunkt ein Anstieg im Auftreten von artikulierten Vokalisationen mit segmentierter Melodiestructur von 20,3% auf 33,6 %.

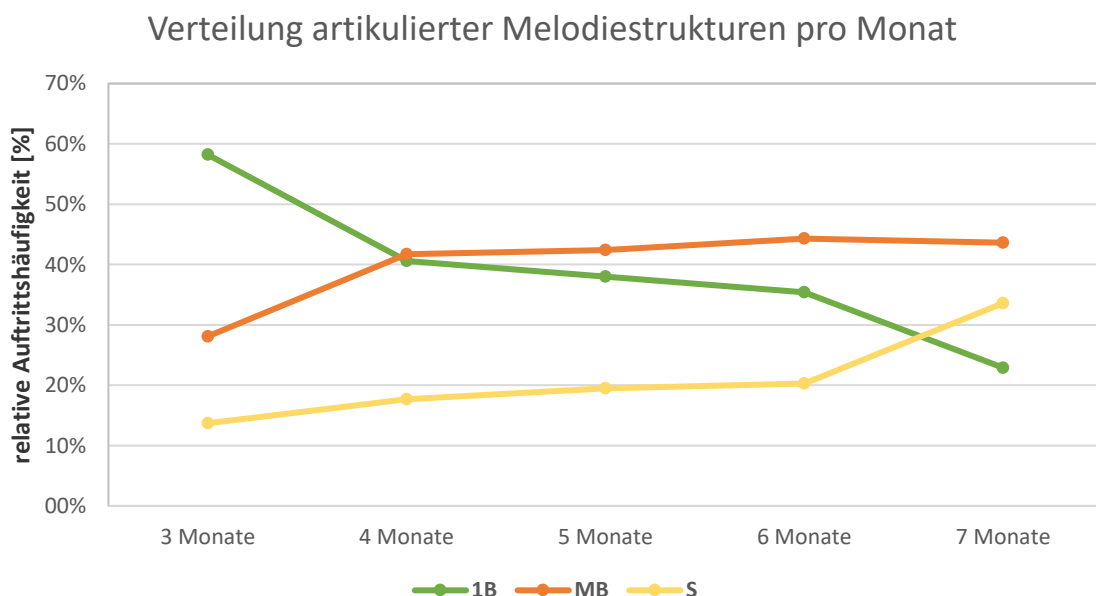


Abbildung 31: Liniendiagramm für die Verteilung artikulierter Vokalisationen über die verschiedenen Melodiestructuren pro Monat (Strukturen mit einem Bogen, Strukturen mit mehreren Bögen, segmentierte Strukturen)

Tabelle 29: Artikulierte Vokalisationen in Abhängigkeit von der Melodiestructur pro Monat (grau hinterlegte Felder)

Alter in Monaten				Lautstruktur			Gesamt	
				1B	MB	S		
3	Artikulation	ja	Anzahl	149	72	35	256	
			%	58,2%	28,1%	13,7%	100,0%	
		nein	Anzahl	265	82	86	433	
			%	61,2%	18,9%	19,9%	100,0%	
	Gesamt			Anzahl	414	154	121	689
				%	60,1%	22,4%	17,6%	100,0%
4	Artikulation	ja	Anzahl	307	315	134	756	
			%	40,6%	41,7%	17,7%	100,0%	
		nein	Anzahl	481	188	176	845	
			%	56,9%	22,2%	20,8%	100,0%	
	Gesamt			Anzahl	788	503	310	1601
				%	49,2%	31,4%	19,4%	100,0%
5	Artikulation	ja	Anzahl	183	204	94	481	
			%	38,0%	42,4%	19,5%	100,0%	
		nein	Anzahl	303	164	131	598	
			%	50,7%	27,4%	21,9%	100,0%	
	Gesamt			Anzahl	486	368	225	1079
				%	45,0%	34,1%	20,9%	100,0%
6	Artikulation	ja	Anzahl	176	220	101	497	
			%	35,4%	44,3%	20,3%	100,0%	
		nein	Anzahl	280	133	118	531	
			%	52,7%	25,0%	22,2%	100,0%	
	Gesamt			Anzahl	456	353	219	1028
				%	44,4%	34,3%	21,3%	100,0%
7	Artikulation	ja	Anzahl	32	61	47	140	
			%	22,9%	43,6%	33,6%	100,0%	
		nein	Anzahl	94	61	37	192	
			%	49,0%	31,8%	19,3%	100,0%	
	Gesamt			Anzahl	126	122	84	332
				%	38,0%	36,7%	25,3%	100,0%
Gesamt	Artikulation	ja	Anzahl	847	872	411	2130	
			%	39,8%	40,9%	19,3%	100,0%	
		nein	Anzahl	1423	628	548	2599	
			%	54,8%	24,2%	21,1%	100,0%	
	Gesamt			Anzahl	2270	1500	959	4729
				%	48,0%	31,7%	20,3%	100,0%

6. Diskussion

In der vorliegenden explorativen Längsschnittstudie wurden erstmalig die Eigenschaften von Komfortvokalisationen von monolingual deutsch aufwachsenden Säuglingen im Alter von 3-7 Monaten quantitativ und qualitativ untersucht. In der internationalen Forschung liegen, abgesehen von Einzelfallstudien, im Wesentlichen nur qualitative Studien vor. Zwar gibt es eine Reihe theoretischer Modelle, die die zeitlichen Abfolgen und charakteristischen Vokalisationen innerhalb aufeinanderfolgender Phasen im ersten Lebensjahr beschreiben, allerdings ohne quantitative Daten, die diese älteren Theorien prüfen.

Die hier erstmals durchgeführte Längsschnittanalyse an einem größeren Datensample hat es ermöglicht, die wesentlichen Aspekte der frühen Entwicklung im laryngealen und supralaryngealen System sowie dem Zusammenspiel beider Systeme objektiv zu charakterisieren. Auf diese Weise wurden erstmalig Referenzwerte für vorsprachliche Komfortvokalisationen deutscher Kinder im Alter von 3-7 Monaten erarbeitet, die für die frühtherapeutische und klinische Anwendung besser als bisher existierende geeignet sein könnten. Zur Charakterisierung der laryngealen Aktivität wurden verschiedene Merkmale der Grundfrequenz, der Komplexität des Melodieverlaufs und der Dauer (Zeit) erfasst. Zur Charakterisierung der supralaryngealen Aktivität wurden die Häufigkeiten und Eigenschaften von konsonantischen Elementen anhand des Artikulationsortes (Artikulationszone) analysiert. Die Interaktion beider Ebenen wurde mittels einer Häufigkeitsanalyse artikulatorischer Elemente in Abhängigkeit von der Melodiestructur erfasst. Im vorliegenden Kapitel erfolgt eine Diskussion und Interpretation der Ergebnisse unter Berücksichtigung des aktuellen Forschungsstandes. Hierbei werden zunächst die Befunde der einzelnen Messgrößen vorgestellt und diskutiert. Parallel dazu werden an geeigneten Stellen Ansatzpunkte für die weiterführende Forschung formuliert.

6.1 Akustische Eigenschaften in Komfortvokalisationen im Alter von 3-7 Monaten

Während das supralaryngeale System im Verlauf der ersten Lebensjahre einen komplexen Reifungsprozess vollzieht, ist das laryngeale System bereits zur Geburt relativ reif und voll funktionsfähig (Bosma, 1975). Die Variation der Grundfrequenz reflektiert die der Lautproduktion zugrundeliegenden respiratorischen und neuronalen Anregungsprozesse auf laryngealer Ebene und stellt somit den Schlüsselparameter bei

der Analyse frühkindlicher Lautproduktionen im Säuglingsalter dar (Wermke, 2002). Auch temporale Eigenschaften spielen eine wichtige Rolle für die rhythmische Einteilung des Phonationsstroms während des Vokalisierens und werden hier analysiert.

6.1.1 Grundfrequenzeigenschaften

Während die Melodiekontur und andere Grundfrequenzeigenschaften von Weinauten im Säuglingsalter bereits relativ gut erforscht sind, liegen zu Grundfrequenzeigenschaften in Komfortvokalisationen von Deutsch erwerbenden Kindern nur sehr wenige Untersuchungen vor (Rothgänger, 2003; Scheiner et al., 2002). Systematische Längsschnittuntersuchungen gab es bisher noch gar nicht. Die Diskussion der eigenen Ergebnisse stützt sich daher vorrangig auf Studien von Kindern anderer Umgebungssprachen. Auch hier liegen vielfach nur Untersuchungen mit wenigen Kindern und/oder zudem sehr heterogenen Probandengruppen vor.

Die Literaturrecherche zu Angaben von Werten für die mittlere Grundfrequenz in Komfortvokalisationen ergab einen Variationsbereich zwischen 290 Hz und 365 Hz (Clement, 2004; Delack, 1976; Delack & Fowlow, 1978; Fuller & Horii, 1986, 1988; Iyer & Oller, 2008; Laufer & Horii, 1977; Murry et al., 1983; Rothgänger, 2003; Ruzza et al., 2003; Scheiner et al., 2002; Wermke et al., 2013). Der Wert der mittleren F_0 (gemittelt über alle Einzelvokalisationen) der vorliegenden Studie ist am oberen Rand dieses Variationsbereiches angesiedelt. Wie in Kapitel 5.1.1 ausführlich dargestellt, lag der Mittelwert von F_0 im Gesamtkorpus der vorliegenden Arbeit bei 361 Hz (Median 323 Hz). Kent und Murray (1982) berichten hingegen in ihrer Querschnittsstudie von deutlich höheren Werten der mittleren F_0 : 445 Hz mit 3 Monaten, 450 Hz mit 6 Monaten und 415 Hz mit 9 Monaten. Die Individualdaten der insgesamt 21 englischsprachig aufwachsenden Säuglinge im Alter von 3, 6 und 9 Monaten lagen dabei zwischen 300 Hz und 500 Hz (Kent & Murray, 1982). Die im Vergleich zu oben aufgeführten Studien relativ hohen Mittelwerte dieser Studie sind wahrscheinlich auf methodische Aspekte zurückzuführen. Die F_0 Analyse heutiger moderner Systeme (z.B. Praat) ist deutlich leistungsfähiger und genauer, stand aber Kent und Murray (1982) noch nicht zur Verfügung. Ein weiterer Einflussfaktor ist möglicherweise die Auswahl der Untersuchungszeitpunkte. So werden zum Zeitpunkt von 3 und 6 Monaten auch von anderen Autoren erhöhte Grundfrequenzwerte berichtet (Clement, 2004; Delack, 1976). Ein weiterer beeinflussender methodischer Aspekt könnte in der Zusammensetzung der Datenkorpora in Bezug auf die in die Untersuchung eingeschlossenen

Vokalisationstypen liegen. Hierzu machen Kent und Murray (1982) jedoch keine genaueren Angaben, außer, dass die ausgewählten Laute den Charakter „speechlike“ hatten und vegetative Laute von der Analyse ausgeschlossen wurden. Die Begriffe „speechlike“ und „non-speechlike“ werden von unterschiedlichen Autoren sehr verschieden unterlegt und verwendet. Kent und Murray (1982) stellen diese Unterteilung selbst in Frage. Die Tatsache, dass sie bei der Interpretation ihrer Ergebnisse auf die Ähnlichkeit zum Schrei hinweisen, lässt vermuten, dass sie bestimmte, in ihren Harmonischen und in ihrer Intensität abgeschwächte, sog. mitigierte Weinlaute, mit in die Analysen eingeschlossen haben. Im ersten Lebensjahr liegt der Variationsbereich der F_0 im Schrei in etwa zwischen 400 und 600 Hz mit einem Mittel von 455 Hz (Michelsson et al., 2002; Rothgänger, 2003; Wermke & Friederici, 2004). Das könnte die höhere mittlere F_0 in den Studien dieser Autoren erklären.

Kent und Murray (1982) interpretieren die Ähnlichkeit zwischen beiden Vokalisationstypen aber nicht allein physiologisch (laryngeale Regulation) sondern postulieren auch eine sprachentwicklungsrelevante Kontinuität aller vorsprachlichen Laute: „This similarity of f_0 values indicates that cry and noncry vocalizations are similar in some aspects of laryngeal regulation and that some degree of continuity might be expected in the mean f_0 frequency between early cry and later cooing and babbling“ (Kent & Murray, 1982, pp. 357–358).

Sowohl die aufgeführten Studien als auch die vorliegende Studie machen deutlich, dass sich die Bereiche der mittleren Grundfrequenz von Komfortvokalisationen und Schreivokalisationen überlappen, die mittlere Tonlage bei Komfortvokalisationen jedoch tendenziell etwas tiefer ist. Das Phonieren in tieferer mittlerer Tonlage ist v.a. auf einen verminderten subglottischen Druck und damit in erster Linie auf physikalischen Ursprung zurückzuführen (Titze, 1994). Verantwortlich sind dafür neurophysiologische Reifungsvorgänge im respiro-laryngealen System, v.a. Verbesserungen in der Atemkontrolle. So ist ein stabiler über längere Lautäußerungen gehaltener subglottaler Luftdruck ab dem dritten Lebensmonat zunehmend besser möglich (Langlois, Baken, & Wilder, 1980). Die Tatsache, dass diese Vokalisationen eher mit geringem Luftdruck produziert werden, könnte auf die Genese neuer entsprechender Reifungsvorgänge im Gehirn zurückzuführen sein. So z.B. die Myelinisierung von Gehirnstrukturen, welche für das Auslösen vokaler Reaktionen und deren Intensitätsmodulation zuständig sind sowie vor allem die Myelinisierung hirnhysiologischer Strukturen welche den kontrollierten Stimmeinsatz sowie deren emotionale Intonation steuern (Deoni et al., 2011; Simonyan & Horwitz, 2011). Diese Entwicklungen führen gleichzeitig zur Fähigkeit einer immer

schnelleren Frequenz- und Intensitätsmodulation. Diese ist nötig um sprachliches Interaktionsverhalten/ Imitationsverhalten (pitch matching) zu entwickeln. So scheint es nicht verwunderlich, dass Säuglinge in diesem Zeitraum auch erstmals die Melodiekontur in Interaktion mit Bezugspersonen imitieren (Lieberman, 1985).

Andererseits scheint es nicht auszureichend zu sein, das Absinken der mittleren Tonlage und die Erweiterung des Vokalisationsrepertoires (auch Imitationsverhalten) ausschließlich auf neurophysiologische Veränderungen im laryngealen System zurückzuführen. Andernfalls sollte sich diese Entwicklungstendenz auch in Schreivokalisationen, welche nach dem dritten Lebensmonat produziert werden, zeigen. Wermke und Kollegen zeigten jedoch in zahlreichen Längsschnittuntersuchungen zu Schreilauten eine relativ stabile mittlere Grundfrequenz ohne Entwicklungstendenzen nach unten (Lind, 1999).

Für die Grundfrequenzveränderungen in den Komfortvokalisationen könnten andererseits auch Entwicklungen auf sozial-kommunikativer und kognitiver Ebene relevant sein. Charakteristisch für die tieferfrequenten Komfortvokalisationen ab dem zweiten Lebensmonat (Coo-Stage) ist der Befund, dass sie emotionale Antworten bei Bezugspartnern motivieren, obwohl sie zunächst weder mit positiven (Lachen) oder negativen (Schrei) Emotionen belegt sind (Oller et al., 2013). So gelten sie in der Regel als neutral und werden häufig in Verbindung mit einem „stirnrunzelnden“ Blick sowie sprech-ähnlichen Mundbewegungen produziert (Trevarthen, 1993). Oller et al. (2013) beschreiben die erstmalige Fähigkeit zur Produktion von Lauten losgelöst von einer unmittelbaren Emotion/Bedürfnis (Protophone) im Alter zwischen drei und vier Monaten als sog. „funktionale Flexibilität“. Sie interpretieren diese „vokale Flexibilität“ als einen wesentlichen Schritt in der menschlichen Sprachevolution bzw. Sprachentwicklung: „[...] the early protophones have a special role in language development and evolution because they are the first sounds to be free of specific fixed functions and thus reveal the opening of a door to the flexibility required in language“ (Oller et al., 2013, p. 6322). Was neben oben aufgeführten anatomischen und neurophysiologischen Reifungsvorgängen den Auslöser für diesen Entwicklungsschritt darstellt ist noch weitgehend unerforscht. Vorliegende Untersuchung postuliert ein angeborenes, universales Entwicklungsprogramm, welches durch sensomotorische und sozial-kognitive Reifungsvorgänge sowie die soziale Interaktion durch nahe Bezugspartner generiert wird. Die klassische Verhaltensforschung und Interaktionsstudien haben gezeigt, dass der menschliche Säugling eine innere Motivation hat, mit seinen Bezugspersonen in Interaktion zu treten. Hierbei richtet sich sein besonderes Interesse v.a. auf das mimische und stimmliche Ausdrucksverhalten des Gegenübers, im

Besonderen, wenn es mit seinem eigenen Ausdrucksverhalten korrespondiert (Papousek, 2014). Deutlich wird dies durch die Nachahmung mimischer und oraler Bewegungsmuster des Gegenübers (Field, Woodson, Greenberg, & Cohen, 1982; Meltzoff & Moore, 1977). „But most of all they seem to have become aware of the impact their cooing, or prebabbling vocal play, has on others and to have begun to make use of it socially to communicate their emotions and demands” (Boysson-Bardies, 1999, p. 39).

Die mittlere Grundfrequenz charakterisiert zwar die Haupttonlage, in der diese Laute produziert werden, für die Beschreibung der vokalen Entwicklung ist es darüber hinaus aber auch wichtig zu beschreiben, welchen innerlautlichen Variationsbereich die Grundfrequenz aufweist und wie sich dieser im Entwicklungszeitraum verändert.

Der in der vorliegenden Studie innerlautlich untersuchte F_0 Range (Hub) liegt gemittelt über den gesamten Untersuchungszeitraum bei 175 Hz (Median von 124 Hz). Laufer und Horii (1977) berichten zwischen dem 4. und 5. Lebensmonat auch Mittelwerte zwischen 161 Hz und 184 Hz (Laufer & Horii, 1977). Delack und Follow (1978) beobachteten im ersten Lebensjahr hingegen Werte zwischen 80 und 110 Hz und liegen somit deutlich unter den hier und in anderen Studien berichteten Werten. Clement (2004) beobachtete mit einem mittleren F_0 -Range von 137 Hz (SD 41Hz) ebenfalls tiefere F_0 -Hub-Werte. Grund dafür ist wahrscheinlich auch hier die Auswahl der untersuchten Vokalisationstypen. Hierzu macht sie allerdings keine genauen Angaben, außer dass vegetative Laute und Schreie von den Analysen ausgeschlossen wurden.

Wie in Kapitel 2 ausführlich beschrieben, fällt der Untersuchungszeitraum hier in eine Phase, in der v.a. der Pitch- und Intensitäts-Range exploriert wird. So werden beispielsweise sehr hochfrequente Vokalisationen (z.B. *squealing*, *yelling*), aber auch Vokalisationen mit sehr tiefen Frequenzen (z.B. *growling*, *going*) produziert (Oller, 1980). Wie in Kapitel 4 beschrieben und Kapitel 3 begründet, erfolgte in der vorliegenden Untersuchung bewusst keine spezielle Auswahl der analysierten Komfortvokalisationen, da das gesamte alterstypische Vokalisationsrepertoire erfasst werden sollte. Dementsprechend gingen in die Berechnung der Grundfrequenzverläufe sowohl sehr hochfrequente Quieker als auch sehr tiefe Brummlaute ein. Dies könnte im Vergleich zu anderen Studien zu einem deutlichen höheren Hub und/oder einer höheren mittleren F_0 führen.

Die altersabhängige charakteristische Erhöhung der laryngealen Variabilität zeigt sich insbesondere im kontinuierlichen Anstieg der Maxima des F_0 -Hubs im Untersuchungszeitraum (siehe **Tabelle 8**), welche aus oben genannten Reifungsvorgängen resultiert.

Das Auftreten alterstypischer Vokalisationsrepertoireerweiterungen spiegelt sich auch im zeitlichen Entwicklungsverlauf der mittleren Grundfrequenz wider. Betrachtet man die Mediane und Mittelwerte der mittleren F_0 auf Basis der Einzelvokalisationen im Untersuchungszeitraum der vorliegenden Studie, so wird deutlich, dass sie sich bis zum 6. Monat zwar relativ stabil verhalten, dass es aber zu einem leichten Anstieg im Median, vor allem aber zu einer größeren interindividuellen Variabilität ab dem 5. Monat kommt (siehe **Tabelle 7**). Betrachtet man die Verteilung der mittleren F_0 -Werte über den Gesamtzeitraum, so wird deutlich, dass diese breiter wird. Die Spannbreite vergrößert sich im 6. Monat um ca. 40% und ist damit deutlich größer als zu anderen Zeitpunkten. Im 7. Monat kommt es wieder zu einem leichten Rückgang. Die kindbasierte Analyse zeigte ebenfalls einen Peak der mittleren F_0 im 6. Monat. Dieser ist jedoch nicht signifikant.

Die in der wenigen und älteren Fachliteratur beschriebenen Befunde zum zeitlichen Entwicklungsverlauf der F_0 -Mittelwerte variieren. Die meisten Autoren konnten – ähnlich wie in vorliegender Studie – keine bzw. nur sehr leichte (nicht signifikante) altersbedingte Veränderungen in diesem Zeitraum feststellen (Delack, 1976; Iyer & Oller, 2008; Kent & Murray, 1982; Laufer & Horii, 1977; Rothgänger, 2003; Scheiner et al., 2002). Delack und Kollegen (1976) beobachteten wie in vorliegender Arbeit auch einen Peak im 6. Monat. Clement (2004) beobachtete als einzige Autorin einen Anstieg der mittleren F_0 mit signifikanten Peaks im Alter von 3,5 und 6 Monaten (Clement, 2004).

Entwicklungsveränderungen beim mittleren F_0 -Range werden hingegen in der Literatur von mehreren Autoren beschrieben und werden durch die Ergebnisse der vorliegenden Studie bestätigt. Hier wird ein kontinuierlicher Anstieg bis zum ersten Lebensjahr beschrieben, wobei es Peaks zwischen dem 4. und 6. Monat gibt (Clement, 2004; Delack, 1976; Laufer & Horii, 1977; Murry et al., 1983). In vorliegender Arbeit konnte auf Grundlage der kindbasierten Analyse ebenfalls ein signifikanter Anstieg vom 3. zum 6. Monat gezeigt werden. Auch der Interquartil-Abstand des F_0 -Ranges (Dynamikbereich der Stimme) vergrößert sich ab dem 5. Monat um mehr als die Hälfte. Hier kommt es ebenfalls im 7. Monat wieder zu einem leichten Abwärtstrend, der jedoch datenbedingt (weniger Kinder) sein kann.

Zusammen genommen deuten die Ergebnisse zum zeitlichen Verlauf der laryngealen Messgrößen auf eine Entwicklung der akustischen Vokalisationseigenschaften im Untersuchungszeitraum hin, die hauptsächlich von der Quelle der Stimmlautproduktion bestimmt sind (Phonationseigenschaften). Besonders deutlich waren Veränderungen

nach dem 4. Monat. Hier zeigten sich eine Vokalisationsrepertoireerweiterung und eine Zunahme der inter-individuellen Variabilität. Diese bilden möglicherweise den Eintritt in die sog. Expansionsphase nach den existierenden vokalen Entwicklungsmodellen ab. Die Zunahme der inter-individuellen Variabilität könnte darauf hinweisen, dass die intentionale Imitation sprachspezifischer Intonationsmuster beginnt.

In der Literatur wird davon berichtet, dass die „Phase der stimmlichen Expansion“ (Oller, 1980; Stark, 1981) oder das „Spiel mit der Stimme“ (Lewis, 1936; Papousek & Papousek, 1981) durch einen hohen Anteil an explorativen Lauten mit hoher Frequenzvariation auffällt. Im Wesentlichen steckt dahinter, dass das resonantorische und artikulatorische Potential des Stimmapparates parallel zur anatomischen und neuromotorischen Reifung (Kapitel 1) erprobt wird, um Kontrolle über hohe und tiefe Stimmregister zu erlangen (Papousek, 1994).

Dies scheint die zeitliche Betrachtung der Phonationseigenschaften der vorliegenden Untersuchung zu bestätigen. Die Forschungsfrage zur Veränderung der stimmphysiologischen Charakteristika in Komfortvokalisationen kann in Bezug auf die Grundfrequenzeigenschaften somit positiv beantwortet werden. Ebenso zeigte sich, dass die methodische Herangehensweise der Autorin bezüglich der Kriterien zur Lautauswahl sinnvoll war (Kapitel 3 und 4). Wären, wie in vielen vorangehenden Studien, bestimmte Vokalisationstypen von den Analysen ausgeschlossen worden, hätte sich die charakteristische Erweiterung des Variationsbereichs der stimmlichen Eigenschaften anhand der Daten nicht abbilden lassen.

Zusammenfassend zeigten die Ergebnisse also einen universalen Trend der hier untersuchten Kinder in Bezug auf die Grundfrequenzeigenschaften im Untersuchungszeitraum. Auffallend war aber auch hier, wie in der Literatur wiederholt beschrieben (z.B. Vihman, 1996), eine teilweise sehr spezifische Entwicklung einzelner Kinder nach dem 4. Monat. Diese spezifische Entwicklung war aufgrund der zu kleinen Stichprobe zwar nicht hauptsächlicher Gegenstand der Arbeit, ist aber zur Interpretation der Ergebnisse, insbesondere vor dem Hintergrund sprachheilpädagogischer Ansätze, relevant. Aus diesem Grund sollen anschließend einige wichtige Befunde der Individualdatenanalyse aufgeführt und diskutiert werden.

So waren z.B. die Maxima der F_0 bei den Kindern sehr unterschiedlich. Diese lagen bei einigen Kindern mehr als doppelt so hoch (z.B. VB, VL, VG) wie bei anderen: Drei von elf Kindern (AC, AD, VH) produzierten beispielsweise keine Laute mit einer mittleren Stimmlage über 1000 Hz. Hierfür könnten unterschiedliche Faktoren verantwortlich sein.

Zum einen könnten methodische Faktoren eine Rolle spielen, zum anderen individuelle Entwicklungsgeschwindigkeiten: Bei dem Kind AC gingen beispielsweise für die Monate 5 und 6, in denen typischerweise hochfrequente *Squeals* und *Quicker* produziert werden, insgesamt leider nur 11 Einzelsignale in die Analyse ein (siehe **Tabelle 2**). Demnach wäre möglich, dass die vorhandene Vokalisationsanzahl nicht das tatsächlich produzierte Vokalisationsrepertoire des Kindes wiedergibt. Vor einer diagnostischen Anwendung in der Zukunft müsste eine größere Mindestanzahl von Vokalisationen pro Kind vorhanden sein und/oder das kindliche Vokalisationsrepertoire anhand geeigneter Eltern-Fragebögen erfasst werden. Bei dem Kind AD gingen im 5. Monat noch genügend Einzelsignale in die Analysen ein, somit lag hier keine methodische Limitation vor. In den Monaten 6 und 7 überwogen bei diesem Kind im Vokalisationsrepertoire allerdings bereits kanonische Babblers, die ausgeschlossen wurden (siehe Kapitel 4). Diese Probandin durchlief die beschriebenen Phasen der vorsprachlichen Entwicklung schneller als der Rest der Gruppe und es kam offenbar bereits im 5. Monat zu einer Reduktion des F_0 -Hubs bzw. einer Fokusverschiebung von sprachlichem Explorationsverhalten zu einem Imitationsverhalten, sodass die Erkundung des Pitchranges hier nicht mehr vorrangig war. Das Kind VH dagegen war möglicherweise noch nicht in der Expansionsphase angekommen.

Die für die hohe Variabilität der Messgrößen zu geringe Stichprobe der vorliegenden Arbeit lässt keine ursächlichen Analysen der interindividuellen Variationen zu. Möglicherweise nutzen Kinder den prinzipiell möglichen F_0 -Range in unterschiedlichem Maße oder erkunden vorrangig den Intensitätsrange oder suprasegmentale Variationsmöglichkeiten. Des Weiteren könnte der Übergang zur Expansionsphase zeitlich individuell sehr variabel sein. Dies müsste an einer größeren Stichprobe und engeren Zeitintervallen untersucht werden. Weiterhin müsste untersucht werden, inwiefern die Anregung durch die Umwelt eine Rolle spielt bzw. sich das Interaktionsverhalten der engen Bezugspersonen auf das produzierte Vokalisationsrepertoire auswirkt. In der vorliegenden Arbeit wurden zumindest erstmalig Referenzwerte für die relevanten Phonationseigenschaften erarbeitet. Weitere systematische Studien an gesunden Kindern und Risikokindern für den Spracherwerb könnten helfen zu entscheiden, ob und wenn ja, in welcher Weise die Grundfrequenzeigenschaften in Komfortvokalisationen geeignet sind, um früher diagnostisch, und damit auch früher als bisher therapeutisch aktiv zu werden.

6.1.2 Die Veränderung der Vokalisationslänge im Untersuchungszeitraum

Die mittlere Vokalisationslänge im Gesamtkorpus der vorliegenden Untersuchung liegt bei $833\text{ms} \pm 670\text{ms}$ und damit im in der Literatur angegebenen Variationsbereich. Dieser ist mit Angaben zwischen 300ms und teils über 1000ms im ersten Lebensjahr allerdings auch sehr groß (Bloom, 1989; Clement, 2004; D'Odorico et al., 1985; Kent & Murray, 1982; Murry et al., 1983; Rothgänger, 2003; Ruzza et al., 2003). Gründe dafür sind auch hier primär wieder unterschiedliche methodische Herangehensweisen. Clement (2004) und Wermke (2013) haben sich nicht auf spezifische Vokalisationstypen beschränkt, ähnlich wie in vorliegender Arbeit, und kamen zu sehr ähnlichen Werten für die mittlere Vokalisationslänge. Wermke und Kollegen (2013) berichten für die deutschen Kinder ihrer Vergleichsstudie zu Komfortvokalisationen von Kindern aus Kamerun (Nso) im dritten Lebensmonat mit einer mittleren Vokalisationslänge von 999ms Werte in einem ähnlichen Bereich wie er in vorliegender Arbeit gemessen wurde (Wermke et al., 2013). Auch Clement (2004) berichtet für normalhörende Kinder ihrer Vergleichsstudie zu Komfortvokalisationen bei normalhörenden Kindern und von Kindern mit Hörschädigung im Alter von 3-12 Monaten mit einer mittleren Vokalisationslänge von 915ms ähnliche Werte. Die etwas höhere mittlere Vokalisationslänge ist in beiden Studien möglicherweise zum einen darauf zurückzuführen, dass auch Vokalisationen der Kategorie „crying-like“ und „whining-like“ (Clement 2004) bzw. sog. Übergangslaute (Wermke et al. 2013) in die Analysen eingeschlossen wurden. Diese mitigierten Übergangsvokalisationen wurden in der vorliegenden Untersuchung ausgeschlossen und sind, da sie eher dem Typ Schreivokalisation zuzuordnen sind, möglicherweise länger. In Forschungsarbeiten zu Schreivokalisationen werden innerhalb der ersten 8 Lebensmonate Lautlängen zwischen 1200ms und 2570ms angegeben (Newman, 2007; Wasz-Höckert, Lind, Vuorenkoski, Partanen, & Valanne, 1968). Der Untersuchungszeitraum bei Clement (2004) erstreckte sich bis zum 12. Lebensmonat. Somit gehen, anders als in vorliegender Arbeit noch größere Lautlängen ab dem 7. Lebensmonat ein, was ebenfalls zu einem größeren Gesamtmittelwert führen könnte. Des Weiteren wurden Kurzlaute ($<300\text{ms}$) im Vergleich zu oben genannten Untersuchungen von vorliegender Analyse nicht ausgeschlossen, da sie über den gesamten Untersuchungszeitraum Bestandteil des Repertoires bleiben. Dieser Befund ist auch bei anderen Autoren zu finden und könnte ein Argument für die Kontinuitätshypothese sein. Nach einer Einübungsphase scheinen diese (kurzen) Elemente anschließend eine Konsolidierung, Differenzierung oder Integration in

komplexere Vokalisationsmuster mit eventuell neuen kommunikativen Funktionen zu erfahren (Kent & Murray, 1982; Oller, 1980; Papousek, 1994; Stark, 1980).

In der Literatur besteht weitgehend Übereinstimmung darin, dass die Vokalisationslänge in Komfortvokalisationen im ersten Lebensjahr kontinuierlich zunimmt (Bettany, 2004; Clement, 2004; D'Odorico et al., 1985; Delack & Fowlow, 1978; Kent & Murray, 1982, 1982; Laufer & Horii, 1977; Murry et al., 1983; Oller, 1980; Rothgänger, 2003). Dabei scheint sich der größte Zuwachs im 4. Monat zu zeigen. So wird in einigen Arbeiten zwischen dem 3. und 4. Monat ein Anstieg der mittleren Lautlänge von bis zu 50% beschrieben (Bettany, 2004; Delack, 1976; Laufer & Horii, 1977; Murry et al., 1983; Oller, 1980).

Auch in vorliegender Arbeit zeigte sich ein kontinuierlicher Anstieg der mittleren Vokalisationslänge bis zum 6. Monat, mit signifikanten Mittelwertunterschieden zwischen Monat 3 und Monat 6. Der leichte Abfall in Monat 7 auf Vokalisationsebene ist wahrscheinlich auf die geringere Anzahl der eingehenden Vokalisationen zurückzuführen. Der Korrelationskoeffizient (Pearson) von $r=0,79$ zwischen Monat 3 und Monat 7 zeigt auch hier einen starken Effekt und lässt vermuten, dass durch eine größere Probandenanzahl hier mit großer Wahrscheinlichkeit signifikante Mittelwertunterschiede aufgedeckt werden könnten. Die Forschungsfrage nach dem Anstieg der mittleren Vokalisationslänge im Untersuchungszeitraum lässt sich somit positiv beantworten. Diese kontinuierliche Zunahme spiegelt sich auch in den individuellen Daten der Kinder wider.

Der Anstieg der mittleren Vokalisationslänge wird im Allgemeinen durch eine zunehmende Reife und Kontrolle über Stimm- und Atmungsmechanismen interpretiert (Oller, 1980). Ein entscheidender Aspekt hierbei ist nach Liebermann (1985) ein inspiratorischer Rückhaltemechanismus, der sog. „hold-back“, der zur Aufrechterhaltung eines kontinuierlichen subglottalen Luftdrucks während der Phonation erforderlich ist. Dieser ist u.a. durch die steilere Rippenstellung beim Neugeborenen noch nicht gegeben und limitiert somit bei Kindern unter 4 Monaten die Länge von Komfortvokalisationen (Lieberman, 1985). Clement (2004) erklärt den Anstieg der mittleren Lautlänge durch einen Entwicklungsschub in der Stimmkontrolle, der die Kinder ab dem Alter von 3,5 Monaten dazu befähigt ihren Stimmeinsatz zeitlich zu kontrollieren. Die zu beobachtende Variation in den Lautlängen ist damit ab diesem Zeitpunkt deutlich größer als zuvor. Dies stimmt mit den Befunden vorliegender Arbeit überein.

Wie die anderen Phonationseigenschaften, zeichnet sich auch die Vokalisationslänge durch eine große interindividuelle Variabilität aus. Die Mediane der einzelnen Kinder liegen zwar relativ nah am Gruppenmedian von 634ms, allerdings unterscheiden sich die einzelnen Kinder stark in der Variationsbreite der mittleren Vokalisationslängen. So ist der Interquartilabstand bei einigen Kindern fast doppelt so groß wie bei anderen. D'Odorico (1984) zeigte in einer Längsschnittstudie zu Schrei- und Nicht-Schrei-Vokalisationen (N= 744) bei vier italienisch-sprachigen Kindern zwischen 4 und 6 Monaten, dass die Lautlänge deutlich vom Melodiemuster beeinflusst wird. Flache melodische Konturen wiesen immer eine geringere Lautlänge auf als steigende und fallende Konturen. „This fact shows the influence of contour on duration, which acts even before babbling“ (D'Odorico, 1984, p. 25).

Auch die Melodiekomplexität nimmt Einfluss auf die mittlere Lautlänge. Beispielsweise untersuchte Bloom (1989) die Lautlänge bei Komfortvokalisationen von 80 Kindern im Alter von 3 Monaten. Äußerungen mit Silbenstruktur und größerer Variabilität in der Melodiestructur waren signifikant länger als vokalische Äußerungen mit einfachen Melodien (Bloom, 1989).

Die Individualdaten der mittleren Vokalisationslänge der vorliegenden Untersuchung liefern ebenfalls Hinweise auf einen möglichen Zusammenhang zwischen Melodiekomplexität und mittlerer Lautlänge. Das Kind AD sticht mit einer mittleren Lautlänge von 1175ms deutlich aus der Probandengruppe heraus und weist gleichzeitig einen sehr hohen Melody Complexity Index von .80 auf (siehe Kapitel 5). Generell lässt sich diese Tendenz auch an den Individualdaten der restlichen Kinder erkennen: je größer der MCI, desto größer die mittlere Vokalisationslänge. Kinder deren Mediane der Lautlänge über dem Gruppenmedian von 634 ms liegen zeigen MCI-Werte zwischen .43 und .80 während Kinder deren Mediane der Lautlänge unter dem Gruppenmedian liegen kleinere MCI-Werte zwischen .36 und .42 aufweisen.

Um diesen Zusammenhang zu beweisen, müsste die Untersuchung an einer größeren Stichprobe durchgeführt werden. Sollte sich gleichzeitig in weiterführenden Untersuchungen der MCI als geeigneter Risikomarker zur Identifikation von Risikokindern erweisen, wäre die mittlere Lautlänge ein schnell und verlässlich erfassbarer Parameter zur Prädiktion von Risikokindern.

6.2 Prosodierelevante Eigenschaften in Komfortvokalisationen im Alter von 3-7 Monaten

Die Melodie wurde als wichtiges prosodisches Element der späteren Sprache bereits in verschiedenen Studien an Säuglingen untersucht. So konnte anhand von Wein-Lauten beispielsweise gezeigt werden, dass sich die Melodiestructur, welche sich objektiv durch den Melodieverlauf und die zeitliche Segmentierung innerhalb einer einzelnen Vokalisation beschreiben lässt, in den ersten Lebensmonaten nach einem angeborenen Entwicklungsprogramm von zunächst einfachen (einböigen) hin zu zunehmend komplexeren (mehrböigen) Strukturen entwickelt (Wermke, 1993, 1994; Wermke et al., 2002; Wermke & Mende, 2011). Die Komplexität dieser Melodiestructuren reflektiert dabei die Modulationskapazität des Säuglings hinsichtlich der Grundfrequenz (F_0). Auf dieser Grundlage konnten Wermke et al. (2007) für den zweiten Lebensmonat einen Risikomarker für die nachfolgende Sprachentwicklung in Form eines Melodiekomplexitätsindex (MCI) postulieren (Wermke et al., 2007).

In vorliegender Arbeit wurde dieser Ansatz erstmals auf Komfortvokalisationen übertragen und angewendet (siehe Kapitel 4). Dabei zeigte sich, dass die Strukturkategorien nach Wermke (2007) auch in Komfortvokalisationen beobachtet werden können. Hier allerdings häufig in tieferer Tonlage und mit kürzerer Einzelvokalisationslänge. Das bedeutet, dass diese Muster nicht durch stimmphysiologische Mechanismen bedingt sind, sondern durch zerebrale Mechanismen generiert werden. Für ihre Expression ist jedoch auch die Reife respiratorisch-laryngealer Mechanismen relevant – hier gezeigt durch die zunehmende Melodielänge mit dem Alter. Dieser Befund der vorliegenden Arbeit ist ein starkes Argument für die Kontinuitätshypothese der sprachlichen Entwicklung. Der Befund, dass sich über den Zeitraum vom 3. -7. Lebensmonat eine Zunahme komplexer melodisch-rhythmischer Strukturen fand, beantwortet die Forschungsfrage vorliegender Arbeit nach der Komplexitätszunahme auf laryngealer Ebene. So zeigte sich eine relativ kontinuierliche Zunahme der Strukturkategorien MB und S und gleichzeitig eine Abnahme der Strukturkategorie 1B, wobei einfache Melodiestructuren im gesamten Zeitraum das Vokalisationsrepertoire dominieren (sinkende Auftrittshäufigkeit von 68% im 3. Monat auf 47% im 7. Monat).

Der Komplexitätsgrad wurde in Anlehnung an die Analyse in Weinlauten mithilfe eines Index objektiv charakterisiert – dem MCI. Bei der kindbasierten monatlichen Analyse zeigte sich eine kontinuierliche Zunahme im Untersuchungszeitraum. Zwischen Monat 3 und Monat 7 konnten signifikante Mittelwertunterschiede aufgedeckt werden ($p=0,02$).

Insbesondere vom 3. zum 5. Monat ($p=0,04$) scheint ein entscheidender Entwicklungssprung stattzufinden. Hier stieg der Median um das Doppelte an. Eine Ausnahme stellt der 6. Monat dar. Der leichte Abfall des MCIs in diesem Monat könnte datenbedingt sein, da hier keine Laute mehr von dem Kind AD eingehen. Diese Probandin brachte in den Monaten vorher die größten MCI-Werte ein (siehe Tabelle 2) und wurde ab dem 6. Monat von der Analyse ausgeschlossen, da sie bereits in der kanonischen Phase angekommen war.

Die Komplexitätszunahme prosodischer Eigenschaften in Komfortvokalisationen im ersten Lebensjahr wird auch in einigen älteren Arbeiten beschrieben. Hsu et al. (2000) untersuchten beispielsweise Komfortvokalisationen von 13 Kindern zwischen dem ersten und sechsten Lebensmonat. Dabei wurden Sprechqualität (syllabic vs. vocalic) und Melodiekomplexität (einfach vs. komplex) unabhängig voneinander perceptiv kodiert. Die Ergebnisse zeigten, dass silbenbildende Vokalisationen länger sind als andere und dass komplexe, rein vokalische Vokalisationen im Gesamtzeitraum weniger oft als andere Vokalisationstypen auftreten (pro Minute). Vor dem vierten Monat werden eher vokalische Lautäußerungen mit einfachen Melodien produziert. Danach treten eher silbenbildende Laute auf, welche komplexe Melodien aufweisen (Hsu et al., 2000). Somit werden die Silben- und Melodiestructur scheinbar parallel mit zunehmendem Alter komplexer. Legerstee (1991) bestätigt, dass zunächst einfache melodische Strukturen im Vokalisationsrepertoire überwiegen und die Anzahl komplexer melodischer Strukturen mit zunehmendem Alter steigt. Des Weiteren konnte sie feststellen, dass die Produktion komplexer melodischer Strukturen in Zusammenhang mit der Mutter-Kind-Interaktion steht. Sie untersuchte längsschnittlich die Vokalisationen von 8 Kindern in Abhängigkeit von unterschiedlichen sozialen Kontexten im Alter von 3 bis 25 Wochen (1-6 Monate). Die Aufnahmen erfolgten im Abstand von 2 Wochen. Zu jedem Aufnahmezeitpunkt wurden den Kindern ihre Mutter, eine für sie fremde weibliche Person und eine Puppe präsentiert, die sowohl aktiv als auch passiv agierten. Es erfolgte eine auditiv-perzeptive Kategorisierung der kindlichen Lautäußerungen in drei Laut-Typen: melodisch (variable Melodiekontur, länger als 500 ms, silbenähnlich, Gurren und Babblen), vokalisches (Serien aus vokal-ähnlichen Lauten, kürzer als 500ms, oft nasal, mit einfacher Melodiekontur), emotional (Lachen, Schreien, mitigierte Laute). Physiologische Laute (Niesen, Schluckauf, Husten, Gurren usw.) wurden von der Analyse ausgeschlossen (Legerstee, 1991). Die Ergebnisse zeigten, dass im Alter von drei und fünf Wochen signifikant weniger melodische Strukturen produziert werden als zu späteren Zeitpunkten. Des Weiteren wurden ab dem Alter von 7 Wochen signifikant mehr melodische Strukturen produziert, wenn die Mutter aktiv in Interaktion mit ihrem

Kind trat. „Although it was difficult to discern audible differences in the vocalizations during the first months of life, by 7 weeks these melodic vocalizations began to be produced primarily in a ‘conversational’ context” (Legerstee, 1991, p. 339). Clement (2004) und Kent & Murray (1982) konnten in ihren Studien ebenfalls bestätigen, dass einfache Melodiestructuren in den ersten Lebensmonaten signifikant häufiger auftreten und dass das Auftreten komplexer Melodiestructuren bis zum sechsten Lebensmonat zunimmt (Clement, 2004; Kent & Murray, 1982). „Thus, from a strictly acoustic perspective, early cry and the later vocalizations of cooing and babbling appear to be vocal performances in continuity “(Kent & Murray, 1982, p. 353).

Offenbar handelt es sich beim Erwerb melodischer Strukturen in vorsprachlichen Vokalisationen um ein universales Entwicklungsprogramm, welches offen für Lernen ist und auf Reifungsvorgängen im Gehirn basiert. Wie in Kapitel 1 ausführlich dargestellt, fallen in den Zeitraum zwischen dem 3. und 5. Monat z.B. ausgeprägte neurophysiologische Reifungsvorgänge, insbesondere die Myelinisierung höherer und an der Vokalisation beteiligter Gehirnregionen. Hierzu zählen der anteriore cinguläre Cortex (ACC) sowie die *Formatio reticularis*. Beim ACC handelt es sich um den vorderen Teil des Gyrus cinguli, welcher sich um den vorderen Teil des Corpus callosum legt. Er ist verantwortlich für einen willkürlich kontrollierten Stimmeinsatz und seine emotionale Intonation. Die *Formatio reticularis* ist u.a. verantwortlich für eine zunehmende Fähigkeit einer schnellen Frequenz- u. Intensitätsmodulation. Beide Gehirnregionen scheinen somit bei der Produktion komplexer Melodiestructuren involviert zu sein und weisen möglicherweise erst ab dem 3. Monat die für die intentionale Imitation von Umgebungslauten notwendige Verarbeitungsgeschwindigkeit auf. Im Alter von drei Monaten sind Kinder bereits in der Lage, die mütterliche Tonlage und Äußerungslänge zu imitieren (Sandner & Wagner, 1981). Demzufolge ist davon auszugehen, dass kindliche Vokalisationen ab dem Alter von drei Monaten bereits intentional produziert werden können und dass die Produktion kindlicher Vokalisationen ab diesem Zeitpunkt zunehmend durch die Eltern-Kind-Interaktion beeinflusst wird.

Neben enormen kortikalen Reifungsvorgängen sind möglicherweise auch Entwicklungen auf der sozial-kommunikativen und kognitiven Ebene für die beschriebene Komplexitätssteigerung prosodierelevanter Eigenschaften in Komfortvokalisationen verantwortlich. Legerstee (1991) zeigte, dass die Produktion komplexer melodischer Strukturen schon sehr früh in Zusammenhang mit der Mutter-Kind-Interaktion steht. Kinder im Alter zwischen 1 und 6 Monaten produzierten in ihrer Längsschnittuntersuchung signifikant mehr komplexe melodische Strukturen, wenn ihre Mütter aktiv in Interaktion mit ihnen traten (Legerstee, 1991). Auch die Qualität von Face-

to-Face Interaktionen, also die Art und Weise wie Bezugspartner verbal und nonverbal mit Kindern in Kontakt treten, nimmt möglicherweise Einfluss auf die Produktion frühkindlicher Vokalisationen. Hsu, Fogel und Messinger (2001) untersuchten Face-to-Face-Interaktionen von 14 Mutter-Kind-Paaren in Bezug auf deren Einfluss auf die Quantität und Qualität kindlicher Komfortvokalisationen (nondistress vocalizations) zwischen 4 und 24 Wochen. Die Ergebnisse ihrer Videoanalysen zeigen, dass sowohl die Quantität (rate per minute) als auch die Qualität (speech-likeness) der Vokalisationen deutlich davon abhängen, ob die Mütter bzw. die Kinder selbst lachen und Blickkontakt besteht: „Infants produced more speech-like syllabic sounds when their mothers were smiling, when they are looking at their mothers' faces, and when the infants themselves were smiling“ (Hsu, Fogel, & Messinger, 2001, p. 107).

Des Weiteren beginnen im Alter von drei Monaten sog. Turn-Taking-Prozesse als Teil der Eltern-Kind-Interaktion, welche ebenfalls Einfluss auf die Qualität kindlicher Vokalisationen haben könnten. Während die Mutter-Kind Interaktionen in den ersten drei Monaten eher als simultanes bzw. zusammenwirkendes verbales Interaktions-Verhalten beschreiben wird, dominiert zum Ende des dritten Monats ein abwechselndes verbales Kommunikationsmuster zwischen Mutter und Kind (Turn-Taking), welches eher dem Dialogverhalten Erwachsener ähnelt (Ginsburg & Kilbourne, 1988; Jasnow & Feldstein, 1986). Berger und Cunningham (1983) zeigten bei Kindern mit Down-Syndrom, dass diese ab 3 Monaten weniger vokalisiert, während ihre Mütter sprachen und mehr vokalisiert, während ihre Mütter schwiegen und indizieren damit den Beginn von Turn-Taking-Verhalten (Berger & Cunningham, 1983). Bloom und Kollegen (1987) untersuchten das Turn-Taking-Verhalten zwischen Kindern und Eltern und dessen Einfluss auf die Qualität kindlicher Vokalisationen. Die Kinder wurden entweder mit dialogorientiertem Turn-Taking oder eher zufälligen Reaktionen ihrer Bezugspartner konfrontiert. Die kindlichen Vokalisationen wurden gezählt und einzeln auditiv kategorisiert: „syllabic“ oder „vocalic“. Die Ergebnisse indizieren einen deutlichen Effekt des Antwortverhaltens der Bezugspersonen auf die Qualität der kindlichen Äußerungen: „When the adult maintained a give-and-take pattern, the infant produced a higher ratio of syllabic/vocalic sounds“ (Bloom, Russell, & Wassenberg, 1987, p. 111). Clement (2004) schlussfolgert, dass das Turn-Taking als Teil der Eltern-Kind-Interaktion die vokale Entwicklung wahrscheinlich ab dem *Expansion Stage* bzw. *Vocal Play* beeinflusst, also ca. ab dem 4. Monat.

Ein typisches Phänomen der Eltern-Kind-Interaktion zeigt sich im sog. *Motherese* oder der *IDS (infant directed speech)*. U.a. wird diese typische Sprechweise durch eine aufmerksamkeitsanregende Prosodie und Intonation sowie durch ein vermindertes

Sprechtempo charakterisiert. Sie hat in erster Linie eine kommunikative und sprachlehrende Funktion und wird als eine spontane Kommunikationsform beschrieben, die nahe Bezugspartner gegenüber Säuglingen und Kleinkindern intuitiv anwenden (Saint-Georges et al., 2013; Zollinger, 2015). Die Präferenz von Kindern für IDS unterliegt im ersten Lebensjahr einem deutlichen Wandel – so lässt sie beispielsweise zum Ende des ersten Lebensjahres deutlich nach. Die Präferenz für IDS ist auch je nach Alter abhängig von ihrer Qualität. So bevorzugen Kinder im Alter von 4 Monaten langsame IDS mit großer Gemütsregung (Affekt), während Kinder im Alter von 7 Monaten eher eine abgeschwächte IDS mit normaler Intonation unabhängig vom Affekt bevorzugen. Kinder im Alter von drei Monaten produzieren mehr vokale Antworten, wenn die IDS eine terminal fallende Kontur, also der englischen Sprachintonation entsprechende Kontur aufweist. Des Weiteren scheint die Präferenz für IDS auch abhängig von der pragmatischen Funktion in Abhängigkeit vom jeweiligen Alter zu sein. Im Alter zwischen 3 und 6 Monaten vollzieht sich ein Sprung von einem eher „beruhigenden“ zu einem eher „zustimmenden“ Charakter und zwischen 6 und 9 Monaten von einem „zustimmenden“ zu einem „leitenden bzw. führenden“ Charakter der IDS (Saint-Georges et al., 2013). Ob die Qualität der kindgerichteten Sprache einen Effekt auf die Produktion von Komfortvokalisationen im ersten Lebensjahr hat, insbesondere auf die Produktion prosodie-relevanter Eigenschaften, ist Gegenstand laufender Forschungsarbeiten.

Auch die Frage danach, inwiefern kognitive Entwicklungen die Produktion früher Vokalisationen beeinflussen, wurde bisher nur in sehr wenigen Studien erforscht. Bloom (1998) postuliert, dass bereits mit drei Monaten ein Zuwachs an kognitiver Aufmerksamkeit verantwortlich für die Vokalisationsentwicklung sein könnte. So können ab diesem Alter beispielsweise intentionale verbale Vokalisationen durch verbale Interaktionen mit Bezugspartnern hervorgerufen werden. Hier sieht sie einen Zusammenhang zu möglichen kognitiven Veränderungen (Bloom, 1998). Davon ausgehend könnten kognitive Leistungen bereits ab dem Eintritt in die Cooing-Phase in enger Beziehung zur Produktion von Komfortvokalisationen stehen und einen deutlichen Zuwachs ihrer Komplexität zwischen dem 3. und 5. Monat erklären.

Diese komplexe Konvergenz an Reifungsvorgängen auf unterschiedlichen Ebenen könnte zum einen die deutlich neue vokale Exploration im Zeitraum zwischen dem 3. und 7. Monat in vorliegender Arbeit erklären, begründet aber wahrscheinlich auch in einem hohen Maße die beobachteten interindividuellen Unterschiede bezüglich der Komplexität untersuchter supra-segmentaler Eigenschaften im Untersuchungszeitraum.

Ob diese Unterschiede auch spracherwerbsrelevant sind, ist bisher noch nicht erforscht und könnte Gegenstand weiterer Arbeiten auf diesem Gebiet sein.

Wie bereits erwähnt, zeigte sich auch in der kindbasierten Analyse eine kontinuierliche Zunahme am Anteil komplexer Melodiestrukturen mit signifikanten Mittelwertunterschieden zwischen Monat 3 und Monat 5 sowie zwischen Monat 3 und 7. Ein deutlicher Entwicklungssprung nach dem 3. Monat spiegelt sich somit auch in der kindbasierten Analyse wider. Zwar verdeutlicht der Anstieg der Gruppenminima, dass die Produktion von einfachen Melodiestrukturen in der Gesamtkohorte abnimmt, zudem da die einzelnen Minima alle von unterschiedlichen Kindern stammen. Dennoch liegen 6 von 11 Kindern unter dem Gruppenmedian von 0,42, während die restlichen 6 Kinder deutlich über diesem Wert liegen. Ein Kind produziert im Mittel sogar doppelt so viele komplexe Melodiestrukturen als andere. Geschlechtsunterschiede können dafür nicht verantwortlich gemacht werden – hier zeigten sich keine signifikanten Unterschiede. Ob bzw. welche oben aufgeführten Reifungsvorgänge und Einflussfaktoren seitens des sprachlichen Inputs der Umgebung für die beobachteten interindividuellen Unterschiede im Output der einzelnen Kinder verantwortlich sind und ob diese auch in Beziehung zu späteren sprachrelevanten Leistungen stehen, müsste in weiterführenden Längsschnittuntersuchungen erforscht werden. Hierfür müssten die Aufnahmesituationen bezüglich des sprachlichen Inputs durch die Bezugspersonen wahrscheinlich noch stärker standardisiert bzw. miterfasst werden und ggf. durch Videoaufnahmen ergänzt werden. Auch die sprachrelevanten Leistungen müssten über einen deutlich längeren Zeitraum mittels standardisierter Testverfahren erfasst werden. Der Befund, dass das Kind mit dem kleinsten MCI-Mittelwert im Untersuchungszeitraum auffällige Ergebnisse im ELFRA-1 (Produktion) mit 12 Monaten sowie im SETK-2 (Produktion II: Sätze) mit 32 Monaten zeigt (**Tabelle 31**), liefert zumindest erste Hinweise auf einen möglicherweise positiven Zusammenhang zu späteren sprachrelevanten Leistungen und damit eine mögliche prognostische Relevanz des MCI's auch in Komfortvokalisationen.

6.3 Artikulatorische Eigenschaften in Komfortvokalisationen im Alter von 3-7 Monaten

Eine weitere Fragestellung der vorliegenden Forschungsarbeit bezieht sich auf die Komplexitätssteigerung in artikulatorisch relevanten Kategorien, insbesondere ob es zu einem Anstieg der Häufigkeit artikulatorischer Elemente in den untersuchten Komfortvokalisationen sowie entsprechend der Reifungsrichtung eine Verschiebung von der hinteren zur vorderen und mittleren Artikulationszone kommt.

Die artikulatorischen Bewegungen innerhalb einer Komfortvokalisation wurden in vorliegender Arbeit mittels spektraler und auditiver Eigenschaften erfasst und systematisch kategorisiert (Kapitel 4). Die Kategorisierung umfasste eine Einteilung in die an der Produktion beteiligten Artikulationszonen. Im Anschluss wurden Häufigkeitsanalysen (laut- und kindbasiert) erstellt. Die Ergebnisse zeigten, dass zum Zeitpunkt des dritten Lebensmonats 30 Prozent der Komfortvokalisationen artikuliert sind. Die Zahl der artikulierten Vokalisationen steigt bis zum 6. Monat auf 43 Prozent an. Auch wenn es zu einem leichten Anstieg in der Produktion konsonantischer Elemente kommt, ist der Großteil der im ersten Halbjahr produzierten Komfortvokalisationen vokalischer Natur.

Betrachtet man die Komplexität der Artikulation, also die Anzahl der artikulatorischen Elemente auf Vokalisationsebene, so zeigt sich eine deutliche Entwicklungstendenz im Untersuchungszeitraum. Während die Produktion von Vokalisationen mit einfacher Artikulation (1 artikulatorisches Element pro Vokalisation) vom 3. bis zum 7. Monat kontinuierlich abnimmt, nimmt die Produktion von Vokalisationen mit komplexer Artikulation (2, 3 und mehr artikulatorische Elemente pro Vokalisation) im Entwicklungsverlauf stark zu. So werden im 7. Monat Laute mit drei oder mehr artikulatorischen Elementen fünfmal öfter als im 3. Monat produziert. Entsprechend wurden Vokalisationen mit einem einzigen artikulatorischen Element mit steigendem Alter weniger häufig produziert.

In Anlehnung an das Modell von Koopmans-van Beinum und Van der Stelt (1986), welches spezifisch auf die Entwicklung von Phonation und Artikulation eingeht, analysierte auch Clement (2004) Quelle (Phonation) und Filter (Artikulation) separat pro Äußerung (Clement, 2004). Dabei werden ähnlich wie in vorliegender Arbeit auch die unterschiedlichen Phonationstypen (Melodiestrukturkategorien) auf laryngealer Ebene kategorisiert. Die Ergebnisse zeigen, ähnlich wie in vorliegender Untersuchung, dass der über den gesamten Zeitraum am häufigsten auftretende Phonationstyp der unsegmentierte Melodieeinzelbogen ohne artikulatorisches Element, also der einfachste

Typ ist. Er kommt mit einer relativen Auftrittshäufigkeit zwischen 40 und 50% vor. In vorliegender Arbeit betrug diese 55%. Diese Befunde verdeutlichen, dass bei Komfortvokalisationen, wie beim Schrei auch, zunächst einfache Strukturen vorherrschen.

Ein markantes Merkmal neben der Entwicklung melodischer Eigenschaften ist die konsonantische Entwicklung und damit die Bildung von Silben. Vor dem zweiten bzw. dritten Lebensmonat ist die Larynxposition noch relativ hoch und Epiglottis und Velum liegen eng aneinander (siehe Kapitel 1). Die Zunge ist relativ groß und füllt die Mundhöhle fast ganz aus. Das Velum und der weiche Gaumen sind relativ schlaff und passiv und berühren fast die Zunge und die Epiglottis (Fletcher, 1973). Demzufolge sind Artikulationsbewegungen im Allgemeinen, besonders aber im zentralen und frontalen Bereich der Mundhöhle nur sehr eingeschränkt möglich. Die meisten Vokalisationen werden mit geschlossenen Lippen produziert. Zunge und Velum liegen noch nah aneinander und Zunge und Kiefer sind im Vergleich zum Erwachsenen noch rückverlagert. Dies äußert sich in vermehrt auftretenden alterstypischen dorsal produzierten Gurrlauten (z.B. velare Frikative und velares Trillern). Zum Zeitpunkt des dritten Monats werden auch in vorliegender Untersuchung noch 40% aller artikulatorischer Elemente im dorsalen Bereich produziert.

Insgesamt zeigte sich bei der Analyse des Artikulationsrepertoires, ähnlich wie bei anderen Autoren auch, in vorliegender Untersuchung entsprechend der Reifungsrichtung eine Entwicklung von „unten“ nach „oben“, also vom Larynx Richtung vordere Artikulationszonen (Benner, Grenon, & Esling, 2007; Irwin, 1946, 1947a, 1947b; Smith & Oller, 1981). Das führt gleichzeitig zu einer Erweiterung des Konsonantenrepertoires. Stoel-Gammon und Otomo (1986) verglichen die Vokalisationen von 11 normal hörenden und 11 hörgeschädigten Kindern im Alter von 4-18 Monaten hinsichtlich der Entwicklung des Konsonantenrepertoires und des Anteils mehrsilbiger CV-Strukturen (Stoel-Gammon & Otomo, 1986). Die Ergebnisse (für die Gruppe der normal hörenden Kinder) zeigten, ähnlich wie in vorliegender Studie einen allgemeinen Anstieg der Anzahl von Konsonanten-Typen mit zunehmendem Alter. So stieg die mittlere Anzahl von Konsonanten-Typen pro Aufnahme von 11,3 (4,0-5,9 Monate) auf 29,5 (16,0-18,4 Monate) an (Stoel-Gammon & Otomo, 1986).

Aus systemtheoretischer Sicht kann man auch hier in gewissem Maße von einer Zunahme der Diversität und Komplexität mit dem Alter sprechen. Während sich bei dorsalen (rückverlagerten) Artikulationen lediglich die hinteren Teile des Zungenrückens in Richtung Rachen, Velum und hinterem Gaumen bewegen, sind bei apikalen und labialen Artikulationen komplexere Bewegungen der Zungenspitze in Richtung vorderer

Gaumen und Schneidezähne sowie Lippenbewegungen erforderlich. Dies führt zu einem Repertoirezuwachs bei Konsonanten und Konsonantenclustern, also zunehmend komplexeren artikulatorischen Bewegungen.

Clement (2004) kommt z.B. zu dem Ergebnis, dass fast 69 Prozent aller artikulatorischen Elemente im Alter zwischen 2,5 und 4,5 Monaten im hinteren Bereich des Vokaltraktes (velar, uvular, pharyngeal) produziert werden. Zum Ende des ersten Lebensjahres sind es nur noch 11 Prozent, während über 50 Prozent der artikulatorischen Elemente im zentralen Bereich des Vokaltraktes (dental, alveolar, palatal) gebildet werden. Die Produktion artikulatorischer Elemente im vorderen Bereich des Vokaltraktes (labial, labio-dental) weist zunächst einen Anstieg auf. So werden im Alter von 6 Monaten über 50% der Artikulationen im frontalen Bereich des Vokaltraktes gebildet (Peak von 67% mit 7,5 Monaten). Danach werden diese von Artikulationen im zentralen Bereich des Vokaltraktes abgelöst. Somit scheint es eine Verschiebung bezüglich des Artikulationsortes im ersten Lebensjahr zu geben: vom hinteren über den vorderen hin zum zentralen Bereich des Vokaltraktes (Clement, 2004). Die Ergebnisse der vorliegenden Arbeit bestätigen diese früheren Untersuchungen.

Im Einzelnen zeigen die Ergebnisse der vorliegenden Untersuchung, dass der Anteil dorsal gebildeter artikulatorischer Elemente zwischen dem 3. und 7. Monat um ca. drei Viertel zurück geht. Der Anteil labial produzierter artikulatorischer Elemente hingegen steigt um zwei Drittel an. Die Häufigkeit der apikal produzierten artikulatorischen Elemente verdoppelt sich sogar. Laryngeal produzierte konsonantische Elemente treten im gesamten Untersuchungszeitraum konstant unter 5% auf. Dieser Befund reflektiert, dass im Untersuchungszeitraum die Erkundung des supralaryngealen Resonanzraums als Artikulator im Vordergrund steht. Dabei scheint ab dem 5. Monat beim Einsatz der Artikulatoren auch das Experimentieren mit Verschlüssen und Engebildungen besonders ausgeprägt zu sein. So nimmt die Produktion von rein vokalen Resonanzveränderungen ohne konsonanten-ähnliche Verschlussbildung ab diesem Zeitpunkt deutlich ab. Die Abnahme könnte jedoch auch partiell methodisch zu begründen sein. Wenn innerhalb einer Vokalisation ein konsonanten-ähnliches Element kategorisiert wurde, wurden aufgrund möglicher co-artikulatorischer Effekte keine vokalen Resonanzveränderungen mehr kategorisiert. Da ab dem 5. Monat die Artikulation in Form von konsonanten-ähnlichen Elementen deutlich komplexer wird, wurden möglicherweise weniger vokale Resonanzveränderungen vergeben.

Verantwortlich für die Verlagerung der Artikulationsbewegungen in frontale und zentrale Bereiche sind auch hier anatomische Entwicklungsveränderungen zwischen dem 4. und 6. Monat sein (siehe Kapitel 1). Der Unterkiefer wächst nach unten. Dadurch hat die

Zunge mehr Raum für Artikulationsbewegungen und berührt nicht zwangsläufig das Velum. Der Larynx senkt sich ab, sodass supra-laryngeale Artikulation und Phonation physikalisch zwei unabhängige, aber fein aufeinander abgestimmte Prozesse während des Vokalisierens werden (Nishimura, Mikami, Suzuki, & Matsuzawa, 2003). Hinzu kommt nach Bloom (1998) wie bereits weiter oben erwähnt eine wachsende neuromotorische Kontrolle der Brustmuskulatur und eine im Verhältnis zur Mundhöhle kleiner werdende Zunge, wodurch mehrere Artikulatoren kontrolliert und dadurch verschiedene Vokalisationstypen korrespondierend zum Vocal Play/Expansion Stage realisiert werden können. Beispielsweise ist es durch einen höheren subglottalen Luftdruck in Kombination mit Schließbewegungen der Lippen möglich, sog. Raspberries zu produzieren (Clement, 2004).

Neben diesen anatomischen Reifungsvorgängen sind auch neurologische Reifungen im Zusammenhang mit dem Beginn und der Dauer der Phasen der vorsprachlichen Entwicklung beschrieben (Netsell, 1981). Hierzu zählt die Myelinisierung des Motorischen Kortex sowie der inter- bzw. intrahemisphärischen Verbindungen im 5. und 6. Monat. Neurologische Störungen sind als Einflussfaktoren auf den Babbling Onset bzw. die Komplexität der Konsonantenproduktion beschrieben (Velleman, 1994). Auch hormonelle Entwicklungen könnten einen Einfluss auf die Hirnentwicklung und damit auf die vorsprachliche Entwicklung, insbesondere die Produktion artikulatorischer Elemente haben. Quast et al. (2016) konnten einen positiven Zusammenhang zwischen der postnatalen Konzentration des Sexualhormons Östrogen zum Zeitpunkt der sog. „Minipubertät“ und den artikulatorischen Leistungen im Alter von 5 Monaten zeigen (Quast, Hesse, Hain, & Wermke, 2016).

Aufgrund der zunehmenden Komplexität der Lautproduktion sowie der möglichen externen Einflüsse werden interindividuelle Einflüsse größer. Möglicherweise verlaufen insbesondere das Längenwachstum des Pharynx sowie das Kieferwachstum interindividuell sehr unterschiedlich. Die kindbasierte Analyse verdeutlichte, dass sich die einzelnen Probanden in der Produktion von artikulatorischen Elementen teilweise stark unterscheiden. Dabei ist die Spannweite für die dorsale Artikulation im Untersuchungszeitraum am größten. Während das Artikulationsrepertoire bei einigen Kindern noch zwischen 40 und 50% aus dorsalen Elementen bestand, zeigten andere diese Häufigkeiten bereits in der Produktion labialer Elemente.

Zusammengefasst lässt sich die formulierte Forschungsfrage der vorliegenden Arbeit nach der Zunahme komplexer Strukturen auch auf supralaryngealer Ebene positiv beantworten: Es erfolgt ein Anstieg der Häufigkeit der Produktion artikulatorischer

Elemente sowie eine Verschiebung von der hinteren zur vorderen und mittleren Artikulationszone im Untersuchungszeitraum.

6.4 Zusammenspiel von Komplexität in Artikulation und Phonation in Komfortvokalisationen im Alter von 3-7 Monaten

In der vorliegenden Arbeit wurde gezeigt, dass die vorsprachliche Entwicklung durch eine Komplexitätszunahme auf laryngealer und supralaryngealer Ebene charakterisiert ist. Damit zeigten sich ähnliche Entwicklungsprozesse wie sie bereits in der Entwicklung von Weinlauten beschrieben wurden: zuerst werden einfache Bausteine bereitgestellt, welche dann stabilisiert und modifiziert werden. Es wurde gezeigt, dass es sich bei Komfortvokalisationen auf laryngealer Ebene um identische Bausteine handelt (Melodiestruktur) – ein weiteres Argument für die Kontinuitätshypothese. Das allgemeine Entwicklungsprinzip von einfach zu komplex ist ein universales Entwicklungsprinzip in der Biologie (Riedl, 1975). Damit zeigt sich, dass biologische Grundlagen sich auch in der vorsprachlichen Entwicklung reflektieren. Zusätzlich werden sie aber stark durch Lernprozesse in der sozial-emotionalen und sprachlichen Umgebung beeinflusst. Dies führt zu interindividuellen Unterschieden, wie sie in den vorangehenden Kapiteln beispielhaft aufgezeigt wurden. Solche komplexen Vorgänge laufen nicht unabhängig voneinander ab, sondern sind ihrerseits wieder in ein größeres System eingebettet. In der vorliegenden Arbeit hat sich dies durch das Zusammenspiel von Melodiekomplexität und Artikulationskomplexität gezeigt.

Betrachtet man die Ergebnisse der Individualdaten für den MCI, so zeigt sich, dass die 4 Probanden (AC, VF, VH, VI) mit einem unterdurchschnittlichen MCI-Wert (unterhalb des Gruppenmedians von .42) geringere artikulatorische Leistungen zeigten als der Rest der Gruppe. Während bei den Kindern AC und VI ein Großteil des Artikulationsrepertoires noch aus rein vokalischen Resonanzveränderungen in Kombination mit wenigen dorsalen Artikulationen besteht, besteht es bei den Kindern VF und VI bereits zu einem Großteil aus dorsaler Artikulation.

Diese Parallelen stützen die Grundannahme der vorliegenden Arbeit, welche postuliert, dass die Melodie in Komfortvokalisationen als „Baugerüst“ für erste artikulatorische Bewegungen fungiert und eine Zunahme der Komplexität auf supralaryngealer Ebene von einer Zunahme der Komplexität auf laryngealer Ebene abhängig ist (Wermke, 2002; Wermke & Mende, 2011). Demzufolge setzt das Experimentieren mit unterschiedlichen

Artikulatoren und das Segmentieren des Phonationsstroms durch artikulatorische Elemente möglicherweise ein bestimmtes Niveau an Melodiekomplexität voraus. Gestützt wird diese Annahme ebenfalls durch die Betrachtung der Ergebnisse der Produktion artikulatorischer Elemente in Abhängigkeit von der Melodiestructur. Hier wird deutlich, dass im 3. Lebensmonat noch fast 60% der artikulierten Laute eine einfache Melodiestructur (1B) aufweisen. Zwischen dem 3. und 4. Lebensmonat nimmt dieser Anteil um fast 20% ab, während der Anteil artikulierter Vokalisationen mit komplexen Strukturen (MB, S) kontinuierlich zunimmt und im 6. Lebensmonat bereits 65% ausmacht. Dies macht deutlich, dass die Entwicklung auf laryngealer und supra-laryngealer Ebene parallel verläuft. Während das Artikulationsrepertoire anfangs zunächst auf Basis einfacher, bereits integrierter Melodiemodule trainiert wird, schalten sich ab dem 4. Monat beide Ebenen zusammen und es erfolgt eine Integration von komplexen Artikulationsbewegungen in komplexe Melodiemodule. In der Folge dieses regelhaften Entwicklungsprogramms kommt es im nächsten Schritt möglicherweise zu einer Einbettung in zielsprachliche Strukturen. Dies äußert sich durch den Beginn der kanonischen Babbelphase. In vorliegender Arbeit zeigt sich vom 6. zum 7. Monat ein deutlicher Anstieg artikulierter Vokalisationen mit segmentierten Melodiestructuren. Dieser Entwicklungssprung könnte den Übergang in die kanonische Phase andeuten, welche für Kinder mit deutscher Muttersprache erstmals systematisch von Pachtner (2016) untersucht wurden.

III Conclusio

Die theoretische Grundlage der vorliegenden Arbeit war die Annahme, dass die vorsprachliche Entwicklung nach einer bestimmten Regelmäßigkeit verläuft, die oft durch verschiedene Phasen beschrieben wird. Die Entwicklungsmuster folgen dabei einem Kontinuitätsprinzip von einfachen hin zu komplexen Strukturen, siehe Kapitel 2 (Koopmans-van Beinum & van der Stelt, 1986; Oller, 1980; Stark, 1980). Wermke und Mende konnten in einigen Untersuchungen, u.a. zur Melodieentwicklung im Säuglingsschrei von monozygoten Zwillingen eindrucksvoll belegen, dass die Entwicklung der Melodie universalen Evolutionsprinzipien, wie dem *Prinzip der modularen Komposition von Komplexität* und dem *Prinzip der Wiederholung und Spezialisierung* folgt (Wermke, 1993, 1994; Wermke, 2002; Wermke & Friederici, 2004; Wermke & Mende, 2009). Die Entwicklung komplexer biologischer Strukturen basiert danach auf dem Prinzip der modularen Komposition aus einfachen Elementen, welche aus angeborenen Vorstufen entstehen, die sich differenzieren und stabilisieren. Phänomenologisch spricht man von Reifungsvorgängen. Die noch unreifen Elemente werden dabei als Protomodule bezeichnet, ausgereifte Elemente als Module. Letztere werden als Bausteine im Gehirn gespeichert und stehen somit zur intentionalen Anwendung in fast jedem Kontext zur Verfügung. Dabei ist die Duplizierung eines Moduls die einfachste Form der Komplexitätssteigerung, während eine höhere Komplexität erst durch die Kombination verschiedener Elemente erreicht wird. Es wird angenommen, dass dieses Modell für die gesamte vorsprachliche Entwicklung gilt (Wermke & Mende, 2011, 2016).

In einigen Studien zur vorsprachlichen Entwicklung (siehe Kapitel 2) wurde die melodische Entwicklung auch in Komfortvokalisationen – ein Vokalisationsverhalten, welches direkt an die frühere Phase, in der Schreivokalisation überwiegen, anknüpft – untersucht. Dabei kamen die Autoren ebenfalls einheitlich zu dem Ergebnis, dass die Komplexität der Melodie innerhalb von Komfortvokalisationen in den ersten sechs Monaten kontinuierlich zunimmt. Vokalisationen mit einfachen Melodien überwiegen in diesem Zeitraum jedoch (Clement, 2004; Hsu et al., 2000; Kent & Murray, 1982; Legerstee, 1991; Oller, 1980; Stark, 1980).

Ziel der vorliegenden Arbeit war es, die Vokalisationen dieser Phase objektiv zu charakterisieren und damit erstmals quantitative Referenzwerte für deutsch-sprachig aufwachsende Kinder hinsichtlich melodischer und artikulatorischer Eigenschaften im

Alter zwischen 3 und 7 Monaten bereit zu stellen. Hauptaugenmerk der spektralanalytischen Untersuchungen lag hierbei auf der Verschaltung der suprasegmentalen und segmentalen Mechanismen: Es sollte die übergeordnete Forschungsfrage beantwortet werden, ob die Melodieentwicklung in Komfortvokalisationen eine Art „Baugerüst“ für erste artikulatorische Bewegungen ist.

Zunächst zeigten die Analysen, dass die Strukturkategorien nach Wermke (2002) auch in Komfortvokalisationen zu beobachten sind. Durch Häufigkeitsanalysen der Melodiestructuren konnte zudem gezeigt werden, dass komplexe Strukturen im Untersuchungszeitraum zunahmten. Dies äußerte sich auch in den Individualdaten der vorliegenden Arbeit: Der Melodiekomplexitätsindex (MCI), welcher im Längsschnitt Aufschluss über den Komplexitätsgrad der untersuchten Komfortvokalisationen gibt, nahm im Untersuchungszeitraum ebenfalls zu. Insbesondere zwischen dem 3. und 5. Monat gab es einen signifikanten Zuwachs. Der Befund vorliegender Arbeit, dass einfache melodische Strukturen in der hier untersuchten Vokalisationsphase jedoch überwiegen, gibt Anlass zu der Annahme, dass das universale Evolutionsprinzip der Wiederholung hinsichtlich der Entwicklung melodischer Strukturen tatsächlich auch für Komfortvokalisationen gilt. Obwohl der Säugling in der Phase des Säuglingsschreis bereits komplexe Melodiemodule erworben hat, scheint er in der darauffolgenden Phase zunächst wieder auf einfache (Proto-) Module zurückzugreifen. Die vorliegende Arbeit lieferte damit erstmalig quantitative Belege für dieses Entwicklungsmodell im Alter von drei bis sieben Monaten.

Ausgehend von den Ergebnissen vorangegangener Arbeiten wird somit auch in vorliegender Arbeit die Annahme bestätigt, dass es im Untersuchungszeitraum zu einer Zunahme komplexer Strukturen auf laryngealer Ebene in Komfort-Vokalisationen kommt. Es wurde belegt, dass die Komplexität der Melodiestructur der Komfort-Vokalisationen kontinuierlich zunimmt.

Die Komplexitätssteigerung innerhalb der prosodisch relevanten Kategorien (Melodie) wurde auch in artikulatorisch relevanten Kategorien gefunden: Es wurden quantitative Daten vorgelegt, die zeigen, dass die Häufigkeit der Produktion artikulatorischer Elemente in Komfort-Vokalisationen ansteigt und gleichzeitig eine Verschiebung von der hinteren zur vorderen und mittleren Artikulationszone erfolgt. Wie in vorangegangenen Arbeiten bereits von anderen Autoren beschrieben, konnte in vorliegender Arbeit durch quantitative Analysen somit belegt werden, dass sich neben melodischen und zeitlichen Eigenschaften im ersten halben Lebensjahr – bedingt durch anatomische und neurophysiologische Reifungsvorgänge auf supralaryngealer Ebene (siehe Kapitel 1) -

auch die Produktion von artikulatorischen Elementen in kindlichen Vokalisationen regelhaft entwickelt (Koopmans-van Beinum & van der Stelt, 1986; Oller, 1980; Stark, 1980). Dabei nimmt zum einen das Verhältnis von artikulierten zu nicht-artikulierten Lauten zu, zum anderen verändern sich die Artikulationszonen im Vokaltrakt (Clement, 2004; Grenon et al., 2007; Irwin, 1946, 1947a, 1947b, 1951; Smith & Oller, 1981; Stoel-Gammon, 1988; Stoel-Gammon & Otomo, 1986).

Des Weiteren fanden sich Belege dafür, dass die Melodie in Komfort-Vokalisationen als „Baugerüst“ für erste artikulatorische Bewegungen fungiert und eine Zunahme der Komplexität auf supralaryngealer Ebene von einer Zunahme der Komplexität auf laryngealer Ebene abhängig ist. Anhand einer Häufigkeitsanalyse der artikulierten Melodiestructuren konnte gezeigt werden, dass die Produktion von artikulatorischen Elementen in Vokalisationen mit einfacher Melodiestructur mit zunehmendem Alter kontinuierlich abnimmt. Hingegen nimmt die Produktion von artikulatorischen Elementen in Vokalisationen mit komplexen Melodiestructuren im Untersuchungszeitraum kontinuierlich zu. Dies macht deutlich, dass die Entwicklung auf laryngealer und supralaryngealer Ebene parallel verläuft. Anfangs wird das Artikulationsrepertoire auf Basis einfacher, bereits integrierter Melodiemodule trainiert. Ab dem 4. Monat schalten sich beide Ebenen zusammen und es folgt eine Integration von komplexen Artikulationsbewegungen in komplexe Melodiemodule. Der nächste Entwicklungssprung könnte sich in Form einer Einbettung in zielsprachliche Strukturen zeigen und sich im Beginn der kanonischen Babbelphase äußern. So zeigte sich in vorliegender Arbeit vom 6. zum 7. Monat ein deutlicher Anstieg artikulierter Vokalisationen mit segmentierten Melodiestructuren. Die Individualdaten zur Charakterisierung der Melodiekomplexität (MCI) sowie der artikulatorischen Leistungen zeigten, dass Kinder mit unterdurchschnittlichen MCI-Werten auch geringere artikulatorische Leistungen produzieren. Diese Parallelen zeigen, dass eine Zunahme der Komplexität auf supralaryngealer Ebene von einer Zunahme der Komplexität auf laryngealer Ebene abhängig ist und beantwortet die Forschungsfrage, ob die Melodie in Komfort-Vokalisationen als „Baugerüst“ für erste artikulatorische Bewegungen fungiert (Wermke, 2002; Wermke & Mende, 2011). Die Ergebnisse der vorliegenden Arbeit verdeutlichen, dass das Experimentieren mit unterschiedlichen Artikulatoren und das Segmentieren des Phonationsstroms durch artikulatorische Elemente mit großer Wahrscheinlichkeit ein bestimmtes Niveau an Melodiekomplexität voraussetzt. Dies müsste allerdings an einer größeren Stichprobe untersucht werden.

Weiterhin konnten in vorliegender Arbeit durch quantitative Analysen Veränderungen der stimmphysiologischen Charakteristika gezeigt werden, welche als mögliche Voraussetzung für die Realisierung dieses Entwicklungsprogramms interpretiert wurden. Die Ergebnisse der Grundfrequenzanalyse in vorliegender Arbeit machten deutlich, dass sich die mittlere Grundfrequenz von Komfortvokalisationen und Schreivokalisationen überlappt, die mittlere Tonlage bei Komfortvokalisationen jedoch tendenziell etwas tiefer ist. Was neben neurophysiologischen und anatomischen Reifungsvorgängen den Auslöser für diesen Entwicklungsschritt – also das Phonieren in tieferer mittlerer Tonlage und damit die Grundlage für die Erweiterung des Vokalisationsrepertoires – verantwortlich ist, ist bisher noch nicht weitreichend untersucht worden. Vorliegende Untersuchung postuliert ein angeborenes, universales Entwicklungsprogramm, welches durch sensomotorische und sozial-kognitive Reifungsvorgänge sowie die soziale Interaktion durch nahe Bezugspartner generiert wird (Wermke & Mende, 2011, 2016). Demzufolge wird das Absinken der mittleren Tonlage und die Erweiterung des Vokalisationsrepertoires (auch Imitationsverhalten) nicht ausschließlich als Folge neurophysiologischer Veränderungen im laryngealen System interpretiert. Andernfalls sollte sich diese Entwicklungstendenz auch in Schreivokalisationen, die nach dem dritten Lebensmonat produziert werden, zeigen. Wermke und Kollegen zeigten jedoch in zahlreichen Längsschnittuntersuchungen zu Schreilauten eine relativ stabile mittlere Grundfrequenz ohne Entwicklungstendenzen nach unten (Lind, 1999).

Charakteristisch für die tieferfrequenten Komfortvokalisationen ab dem zweiten Lebensmonat (Coo-Stage) ist der Befund, dass sie emotionale Antworten bei Bezugspartnern motivieren (Oller et al., 2013). Umgekehrt könnten diese Reaktionen den Säugling dazu motivieren, sprach-ähnliche Vokalisationen im Rahmen seiner anatomischen und neurophysiologischen Möglichkeiten zu produzieren, um in kommunikativen Austausch mit seinen Bezugspartnern zu treten. Demzufolge wäre es möglich, dass dieses angeborene Entwicklungsprogramm schon sehr früh durch soziale Interaktionen beeinflusst wird und damit auch offen für eine sprachheilpädagogische Förderung von außen ist - z.B. durch Interaktionstraining der nahen Bezugspartner. Hierfür müsste in weiteren Forschungsarbeiten an einer größeren Stichprobe zunächst jedoch die Rolle der Umwelt in der frühen Vokalisationsentwicklung systematisch untersucht werden. Methodisch könnten sich hier neben Audio-Aufnahmen auch Videoanalysen und Elternbefragungen eignen.

Neben der mittleren Tonlage war für die Beschreibung der vokalen Entwicklung darüber hinaus aber auch wichtig zu beschreiben, welchen innerlautlichen Variationsbereich die Grundfrequenz aufweist und wie sich dieser im Entwicklungszeitraum verändert. In

vorliegender Arbeit konnte auf Grundlage der kindbasierten Analyse des F_0 -Ranges ein signifikanter Anstieg vom 3. zum 6. Monat gezeigt werden. Auch der Interquartil-Abstand des F_0 -Ranges vergrößert sich ab dem 5. Monat um mehr als die Hälfte. Hier zeigte sich die altersabhängige charakteristische Erhöhung der laryngealen Variabilität insbesondere im kontinuierlichen Anstieg der Maxima des F_0 -Hubs, welche aus oben genannten Reifungsvorgängen resultiert. Neben der Vergrößerung des stimmlichen Dynamikbereichs zeigte sich ebenfalls eine Zunahme der inter-individuellen Variabilität nach dem 4. Monat. Diese Befunde stellen möglicherweise einen Beleg für den Eintritt in die sog. Expansionsphase nach den existierenden vokalen Entwicklungsmodellen dar. Die Zunahme der inter-individuellen Variabilität könnte ein Hinweis darauf sein, dass die intentionale Imitation sprachspezifischer Intonationsmuster beginnt.

Im Wesentlichen konnten die Daten der phonatorischen und artikulatorischen Eigenschaften zeigen, dass im Untersuchungszeitraum das resonatorische und artikulatorische Potential des Stimmapparates erprobt wird. Dies ist v.a. an anatomische und neuromotorische Reifungsprozesse geknüpft, welche das Kind nach und nach dazu befähigen, Kontrolle über hohe und tiefe Stimmregister zu erlangen und artikulatorische Bewegungen zu koordinieren (Papousek, 1994).

Inwiefern die Anregung durch die Umwelt in der Ausprägung der stimmphysiologischen und artikulatorischen Charakteristika in Komfortvokalisationen eine Rolle spielt, muss in zukünftigen Projekten weiter untersucht werden.

Neben der melodischen und artikulatorischen Komplexität kristallisierte sich in vorliegender Arbeit v.a. die mittlere Vokalisationslänge als vielversprechender Untersuchungsparameter für weitere Forschungsprojekte. In Übereinstimmung mit vorangegangenen Arbeiten konnte auch in vorliegender Arbeit gezeigt werden, dass die mittlere Vokalisationslänge vom 3. bis zum 6. Monat kontinuierlich zunimmt. Zwischen Monat und 3 und Monat 6 konnten signifikante Mittelwertunterschiede aufgedeckt werden. Einige Autoren beschreiben den größten Zuwachs im 4. Monat (Bettany, 2004; Delack, 1976; Laufer & Horii, 1977; Murry et al., 1983; Oller, 1980). Dieser wird im Allgemeinen durch eine zunehmende Reife und Kontrolle über Stimm- und Atmungsmechanismen interpretiert (Oller, 1980). Dieser Entwicklungsschub in der Stimmkontrolle befähigt die Kinder ab einem Alter von 3,5 Monaten wahrscheinlich dazu, ihren Stimmeinsatz zeitlich zu kontrollieren. Die zu beobachtende Variation in den Lautlängen ist damit ab diesem Zeitpunkt deutlich größer als zuvor (Clement, 2004). Dies stimmt mit den Befunden vorliegender Arbeit überein. Ein weiterer Befund vorliegender Arbeit zeigte sich in der Betrachtung der Individualdaten der mittleren Vokalisationslänge. Hier zeigten sich Hinweise auf einen potenziellen Zusammenhang

zur Melodiekomplexität in den untersuchten Komfortvokalisationen: Kinder mit größerem Melodiekomplexitätsindex (MCI) zeigten größere mittlere Vokalisationslängen. Kinder deren Mediane der Lautlänge über dem Gruppenmedian von 634 ms liegen zeigen MCI-Werte zwischen .43 und .80 während Kinder deren Mediane der Lautlänge unter dem Gruppenmedian liegen kleinere MCI-Werte zwischen .36 und .42 aufweisen. Um diesen Zusammenhang zu beweisen, müsste die Untersuchung an einer größeren Stichprobe wiederholt werden. Sollte sich gleichzeitig in weiterführenden Untersuchungen der MCI als geeigneter Risikomarker zur Identifikation von Risikokindern für spätere Sprachprobleme erweisen, wäre die mittlere Lautlänge ein schnell und verlässlich erfassbarer Parameter zur Prädiktion von Risikokindern.

Bloom (1989) konnte zeigen, dass Kinder im Altern von 3 Monaten signifikant größere Vokalisationslängen zeigen, wenn Sie Äußerungen mit Silbenstruktur, also artikulatorischen Elementen, produzieren. Gleiches gilt für Äußerungen mit größerer Variabilität in der Melodiestructur (Bloom, 1989). Beide in vorliegender Arbeit untersuchten Parameter – Melodie und Artikulation – stehen offensichtlich in engem Zusammenhang mit der produzierten Vokalisationslänge in Komfortvokalisationen. Sollte sich in zukünftigen Längsschnittprojekten bestätigen, dass die drei Parameter in Zusammenhang mit späteren sprachlichen Leistungen stehen, könnte dies einen wichtigen Schritt für die Erfassung von Risikomarkern in Bezug auf spätere sprachliche Leistungen darstellen. So könnte möglicherweise die Vokalisationslänge als diagnostisches Kriterium schnell und einfach im Rahmen von Früherkennungsuntersuchungen erfasst werden. Einen weiteren Einsatzbereich könnte die (vor-)sprachliche Verlaufskontrolle während der Hörgeräteanpassung von hörgeschädigten Kindern im ersten Lebensjahr darstellen. Durch die objektive Charakterisierung akustisch-melodischer und artikulatorischer Eigenschaften der Komfortvokalisationen könnte der individuelle Benefit der Versorgung durch Hörgeräte oder CI-Implantate objektiv beurteilt werden und in einem weiteren Schritt Anpassungsprozesse optimiert werden. In einigen Arbeiten wurde bereits dargestellt, dass der individuelle Hörstatus schon in den ersten 6 Lebensmonaten, also schon früher als lange Zeit angenommen, Einfluss auf die Qualität vorsprachlicher Komfortvokalisationen hat. So wurden beispielsweise Unterschiede in der mittleren Grundfrequenz, der Melodiekomplexität und der artikulatorischen Aktivität zwischen hörgeschädigten und normal hörenden Kindern in den ersten 6 Lebensmonaten aufgezeigt (Blum, 2020; Shehata-Dieler et al., 2013; Clement, 2004).

Die vorliegende Arbeit liefert mit entsprechenden Referenzwerten für gesunde Kinder einen wichtigen Beitrag als Vergleichsgrundlage für objektive Kriterien zur Beurteilung

(vor-)sprachlicher Leistungen von Risikokindern. Inwiefern sich die hier untersuchten melodischen und artikulatorischen Eigenschaften tatsächlich als Prädiktoren für spätere sprachliche Leistungen eignen, muss in zukünftigen Längsschnittprojekten an größeren Stichproben untersucht werden. Auch die Rolle des sprachlichen Inputs durch die Umwelt und damit die Frage danach, ob ein frühes präventives Eingreifen in Form sprachheilpädagogischer Maßnahmen einen ungestörten Ablauf der vorsprachlichen und damit der späteren sprachlichen Entwicklung ermöglicht, könnte durch zukünftige Forschungsbemühungen beantwortet werden. Die vorliegenden Ergebnisse liefern jedoch erste Hinweise darauf, dass individuelle Unterschiede nicht allein auf Reifungsparameter zurückzuführen sind.

Literaturverzeichnis

- Bartl-Pokorny, K. D., Marschik, P. B., Sachse, S., Green, V. A., Zhang, D., van der Meer, L., Einspieler, C. (2013). Tracking development from early speech-language acquisition to reading skills at age 13. *Dev. Neurorehabil.*, 16(3), 188–195.
- Benner, A., Grenon, I., & Esling, J. (2007). *Infants' Phonetic Acquisition of Voice Quality Parameters in The First Year of Life* (Proceedings of the 16 th International Congress Phonetic Sciences). Saarbrücken, pp. 2073–2076.
- Berger, J., & Cunnigham, C. (1883). Early social interactions between infants with Down's syndrome and their parents. *Health Visitors*, 56(2), 58–60.
- Bettany, L. (2004). *Range Exploration of Phonation and Pitch in the First Six Months of Life* (Master Thesis). University of Victoria, Victoria.
- Bloom, K. (1989). Duration of Early Vocal Sounds. *Infant Behaviour and Development*, 12, 245–250.
- Bloom, K. (1998). *The missing link's missing link: Syllabic vocalizations at 3 months of age: Kommentar zu Mac Neilage (1998): The frame/content theory of evolution of speech production* (No. 4). *Behavioral and Brain Sciences*, 21, pp. 514–515.
- Bloom, K., Russell, A., & Wassenberg, K. (1987). Turn taking affects the quality of infant vocalizations. *Journal of Child Language*, 14(2), 211–227.
- Blum, Daria Maria (2020): Vergleichende Analyse melodischer Eigenschaften von frühen Komfortlauten bei gesunden und hochgradig hörgeschädigten Säuglingen. Doctoralthesis. Universität Würzburg. Online verfügbar unter <https://opus.bibliothek.uni-wuerzburg.de/frontdoor/index/index/docId/21139>.
- Boersma, P., & Weenink, D. (2009). Doing phonetics by computer: Version 5.1.05. Retrieved from www.praat.org
- Bohland, J. W., & Günther, F. H. (2006). An fMRI investigation of syllable sequence production. *Neuroimage*, 32, 821–841.
- Boseley, M., & Hartnick, C. J. (2006). Development of the Human True Vocal Fold: Depth of Cell Layers and Quantifying Cell Types Within the Lamina Propria. *Annals of Otolaryngology, Rhinology & Laryngology*, 115(10), 784–788.

- Bosma, J. F. (1975). Anatomic and Physiologic Development of the Speech Apparatus. In D.B. Tower (Ed.), *The Nervous System: Human Communication And Its Disorders. Volume 3* (Vol. 3, pp. 469–481). New York: Raven.
- Boysson-Bardies, B. de. (1999). *How language comes to children: From birth to two years*. Cambridge: MIT Press. Retrieved from <http://www.worldcat.org/oclc/38853989>
- Brauer, J., Anwander, A., Perani, D., & Friederici A. D. (2013). Dorsal and ventral pathways in language development. *Brain and language*, 127(2), 289–295.
- Buschmann, Anke (2017): Heidelberger Elterntermin fröhe Sprachförderung. HET Late Talkers. 3. Auflage. München: Urban & Fischer in Elsevier. Online verfügbar unter <http://shop.elsevier.de/978-3-437-44497-5>.
- Cierpka, M. (Ed.). (2014). *Fröhe Kindheit 0-3 Jahre: Beratung und Psychotherapie für Eltern mit Säuglingen und Kleinkindern* (2nd ed.). Heidelberg: Springer.
- Clement, C. J. (2004). *Development of vocalizations in deaf and normally hearing infants* (Dissertation). Universiteit van Amsterdam, Amsterdam.
- D’Odorico, L. (1984). Non-segmental features in prelinguistic communications: An analysis of some types of infant cry and non-cry vocalizations. *Journal of Child Language*, 11(1), 17–27.
- D’Odorico, L., & Franco, F. (1991). Selective production of vocalization types in different communication contexts. *Journal of Child Language*, 18(3), 475–499.
- D’Odorico, L., Franco, F., & Vidotto, G. (1985). Temporal Characteristics In Infant Cry And Non-Cry Vocalizations. *Language and Speech*, 28(1), 29–46.
- Dannenbauer, F. M. (2009). Spezifische Sprachentwicklungsstörung. In M. Grohnfeldt (Ed.), *Lehrbuch der Sprachheilpädagogik und Logopädie* (3rd ed.). Stuttgart: Kohlhammer.
- Delack, J. B. (1976). Aspects of infant speech development in the first year of life. *The Canadian Journal of Linguistics*, 21(1).
- Delack, J. B., & Fowlow, P. J. (1978). The Ontogenesis of Differential Vocalization: Development of Prosodic Contrastivity During the First Year of life. In N. Waterson &

- C. E. Snow (Eds.), *The development of communication* (pp. 93–110). New York: John Wiley & Sons.
- Deoni, S., Mercure, E., Blasi, A., Gasston, D., Thomson, A., Johnson, M., Murphy, D. (2011). Mapping Infant Brain Myelination with Magnetic Resonance Imaging. *The Journal of Neuroscience*, 31(2), 784–791.
- Dubois, J., Hertz-Pannier, L., Dehaene-Lambertz, G., Cointepas, Y., & Le Bihan, D. (2006). Assessment of the early organization and maturation of infants' cerebral white matter fiber bundles: a feasibility study using quantitative diffusion tensor imaging and tractography. *Neuroimage*, 30(4), 1121–1132.
- Eckel, H.E., Sprinzi, G.M., Sittel, C., Koebke, J., Damm, M., & Stennert, E. (2000). Zur Anatomie von Glottis und Subglottis beim kindlichen Kehlkopf. *HNO*, 48(7), 501–507. Retrieved from <http://www.springerlink.com/content/xjlk7nnjvtvte1dg8/fulltext.pdf>
- Egert, F., Sachse, S., & Groth, K. (2016). Title Registration for a Systematic Review: Language Interventions for Improving the L1 and L2 Development of Dual Language Learners in Early Education and Care: A Systematic Review and Meta-Analysis. *Technical Report*. Retrieved from DOI: 10.13140/RG.2.2.27650.0736
- Fasolo, M., Majorano, M., & D'Odorico, L. (2008). Babbling and first words in children with slow expressive development. *Clinical linguistics & phonetics*, 22(2), 83–94.
- Field, T. R., Woodson, R., Greenberg, R., & Cohen, D. (1982). Discrimination and imitation of facial expressions by neonates. *Science*. (218), 179–181.
- Fletcher, S. G. (1973). Maturation of the Speech Mechanism. *Folia Phoniatica*, 25, 161–172.
- Fried, M. P., Kelly J. H., & Strome, M. (1982). Comparison of the Adult and Infant Larynx. *The Journal of Family Practice*, 15(3), 557–561.
- Fuller, B.F., & Horii, Y. (1986). Differences in fundamental frequency, jitter, and shimmer among four types of infant vocalizations. *Journal of Communication Disorders*, 19(6), 441–447.
- Fuller, B.F., & Horii, Y. (1988). Spectral Energy Distribution In Four Types Of Infant Vocalizations. *Journal of Communication Disorders*, 21, 251–261.

- Ginsburg, G. P., & Kilbourne, B. K. (1988). Emergence of vocal alternation in mother-infant interchanges. *Journal of Child Language*, *15*, 221–235.
- Gratier, M., & Devouche, E. (2011). Imitation and repetition of prosodic contour in vocal interaction at 3 months. *Developmental Psychology*, *47*(1), 67–76. <https://doi.org/10.1037/a0020722>
- Gratier, M., Devouche, E., Guellai, B., Infanti, R., Yilmaz, E., & Parlato-Oliviera, E. (2015). Early development of turn-taking in vocal interaction between mothers and infants. *Frontiers in Psychology*, *6*(1167).
- Gray, S. D., Hirano, M., & Sato, K. (1993). Molecular and Cellular Structure of Vocal Fold Tissue. In I. R. Titze (Ed.), *Vocal fold physiology: Frontiers in basic science* (pp. 1–35). San Diego, Calif: Singular Pub. Group.
- Grenon, I., Benner, A., & Esling, J. (2007). *Language-Specific Phonetic Production Patterns In The First Year Of Life* (Proceedings of the 16 th International Congress Phonetic Sciences). Saarbrücken, pp. 1561–1564.
- Grimm, H., & Doil, H. (2000). *ELFRA: Elternfragebögen für die Früherkennung von Risikokindern*. Handanweisung. Göttingen: Hogrefe.
- Grimm, H., & Weinert, S. (2002). Sprachentwicklung. In R. Oerter & L. Montada (Eds.), *Entwicklungspsychologie* (pp. 517–550). Weinheim: Beltz.
- Grossheinrich, N., Schulte-Körne, G., Marschik, P., Kademann, S., Suchodoletz, W. von, & Sachse, S. (2019). School-age outcomes of late-talking toddlers: Long-term effects of an early lexical deficit. *Developmental Science*. (e12826), 1–13. Retrieved from <https://doi.org/10.1111/desc.12826>
- Harding, R. (1984). Function of the Larynx in the Fetus and Newborn. *Annual Review Of Physiology*. (46), 645–659.
- Hartnick, C. J., Rehbar, R., & Prasad, V. (2005). Development of Maturation of the Pediatric Human Vocal Fold Lamina Propria. *The Laryngoscope*, *115*(15), 4–15.
- Hsu, H.-C., Fogel, A., & Cooper, R. (2000). Infant Vocal Development during the First 6 Months: Speech Quality and Melodic Complexity. *Infant and Child Development*, *9*, 1–16.

- Hsu, H.-C., Fogel, A., & Messinger, D. (2001). Infant non-distress vocalization during mother-infant face-to-face interaction: Factors associated with quantitative and qualitative differences. *Infant Behavior & Development*, *24*, 107–128.
- Ibbotson, P., & Tomasello, M. (2017). Wie Sprache entsteht: Ein neues Bild der Sprache. *Spektrum der Wissenschaft*. (3), 12–17.
- Irwin, O. C. (1946). Development of speech during infancy: curve of phonemic types. *Journal of Experimental Psychology: Human Perception & Performance*, *36*(5), 431–436.
- Irwin, O. C. (1947a). Infant Speech: Consonant Sounds According To Manner Of Articulation. *Journal of Speech Disorders*, *12*, 402–404.
- Irwin, O. C. (1947b). Infant Speech: Consonantal Sounds According To Place of Articulation. *Journal of Speech Disorders*, *12*, 397–401.
- Irwin, O. C. (1951). Infant Speech: Consonantal Position. *Journal of Speech and Hearing Disorders*, *16*, 159–161.
- Iyer, S. N., & Oller, D. K. (2008). Fundamental frequency development in typically developing infants and infants with severe-to-profound hearing loss. *Clinical linguistics & phonetics*, *22*(12).
- Jasnow, M., & Feldstein, S. (1986). Adult-like temporal structure of face-to-face communication between mothers and infants. *Child development*, *57*, 754–761.
- Jürgens, U. (2002). Neural pathways underlying vocal control. *Neuroscience and Biobehavioral Reviews*, *26*, 235–258.
- Kent, R. D. (1976). Anatomical and neuromuscular maturation of the speech mechanism: evidence from acoustic studies. *Journal of Speech & Hearing Research*, *19*(3), 421–447.
- Kent, R. D. (1984). Psychobiology of speech development: coemergence of language and a movement system. *The American Journal of Physiology*.
- Kent, R. D. (1995). Development of the craniofacial-oral-laryngeal anatomy: A review. *Journal of Medical Speech-Language Pathology*. (3), 145–190.

- Kent, R. D., & Miolo, G. (1996). Phonetic Abilities in the First Year of Life. In P. Fletcher & B. MacWhinney (Eds.), *The Handbook of child language*. Oxford, UK, Cambridge, Mass., USA: Blackwell.
- Kent, R. D., & Murray, A. D. (1982). Acoustic Features of Infant Vocalic Utterances at 3, 6, and 9 Months. *Journal of the Acoustical Society of America*, 72(2), 353–365.
- Kent, R. D., & Vorperian, H. K. (1995). *Development of the craniofacial-oral-laryngeal anatomy*. San Diego: Singular Pub. Group. Retrieved from <http://www.worldcat.org/oclc/32853109&ap=citavi>
- Koopmans-van Beinum, F. J., Clement, C. J., & Van den Dikkenberg-Pot, C. J. (2001). Babbling and the lack of auditory speech perception: a matter of coordination? *Developmental Science*, 4(1), 61–70.
- Koopmans-van Beinum, F. J., & van der Stelt, J. M. (1986). Early Stages in the Development of Speech Movements. In B. Lindblom & R. Zetterström (Eds.), *Precursors of early speech: Wenner-Gren International Symposium Series. Proceedings of an International Symposium held at the Wenner-Gren Center, Stockholm* (Vol. 44, pp. 37–54).
- Langlois, A., Baken, R. J., & Wilder, C. N. (1980). Pres-speech respiratory behaviour during the first year of life. In T. Murry & J. Murry (Eds.), *Infant Communication: Cry and Early Speech*. Houston, Texas: College-Hill Press.
- Laufer, M. Z., & Horii, Y. (1977). Fundamental frequency characteristics of infant non-distress vocalization during first twenty-four weeks. *Journal of Child Language*, 4, 171–184.
- Legerstee, M. (1990). Infants use multimodal information to imitate speech sounds. *Infant Behavior & Development*, 13(3), 343–354. Retrieved from [https://doi.org/10.1016/0163-6383\(90\)90039-B](https://doi.org/10.1016/0163-6383(90)90039-B)
- Legerstee, M. (1991). Changes in the quality of infant sounds as a function of social and nonsocial stimulation. *First Language*, 11, 327–343.
- Leroy, F., Glasel, H., Dubois, J., Hertz-Pannier, L., Thirion, B., Mangin, J. F., & Dehaene-Lambertz, G. (2011). Early maturation of the linguistic dorsal pathway in human infants. *The Journal of Neuroscience*, 31(4), 1500–1506.

- Lester, B. M. (Ed.). (1985). *Infant Crying: Theoretical and Research Perspectives* (1st ed.). Boston: Plenum Press.
- Lieberman, P. (1967). *Intonation, perception, and language*. Cambridge, MA: MIT Press.
- Lieberman, P. (1985). The Physiology of Cry and Speech in Relation to Linguistic Behavior. In B. M. Lester (Ed.), *Infant Crying: Theoretical and Research Perspectives* (1st ed., pp. 29–54). Boston: Plenum Press.
- Lieberman D.E., McCarthy R.C., Hiemae K.M., & Palmer J.B. (2001). Ontogeny of postnatal hyoid and larynx descent in humans. *Archives Of Oral Biology*. (46), 117–128.
- Lind, K. (1999). *Ontogenetische Entwicklung vorsprachlicher Lautäußerungen in den ersten 3 Lebensmonaten: Diplomarbeit, Humboldt Universität zu Berlin*.
- Loucks T. M., Poletto C. J., Simonyan K., Reynolds C. L., & Ludlow C. L. (2007). Human Brain Activation during Phonation and Exhalation: Common Volitional Control for Two Upper Airway functions. *Neuroimage*, 36(1), 131–143.
- Luthe, L., Hausler, U., & Jürgens, U. (2000). Neuronal activity in the medulla oblongata during vocalization. A single-unit recording study in the squirrel monkey. *Behavioral and Brain Sciences*. (116), 197–210.
- Masataka, N. (2003). *The onset of language*. Cambridge, UK, New York: Cambridge University Press.
- Meltzoff, A. N., & Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science*. (198), 75–78.
- Meyer, T. K. (2009). The Larynx for Neurologists. *The Neurologist*. (6), 313–318.
- Michelsson, K., Eklund, K., Leppänen, P., & Lyytinen, H. (2002). Cry characteristics of 172 healthy 1-to-7-day old infants. *Folia Phoniatica*, 54(4), 190–200.
- Monnier, M., & Willi, H. (1953). The integrative activity of the nervous system of a mesorhombencephalic anencephalus.II. Anatomical part. *Monatsschrift Psychiatr Neurol*. (126), 225–273.

- Murry, T., Hoit-Dalgaard, J., & Gracco, V. L. (1983). Infant Vocalization: A Longitudinal Study of Acoustic and Temporal Parameters. *Folia Phoniatica*, 35(5), 245–253.
- Nathani, S., Ertmer, D., & Stark, R. E. (2006). Assessing vocal development in infants and toddlers. *Clinical linguistics & phonetics*, pp. 351–369.
- Netsell, R. (1981). The Acquisition of Speech Motor Control: A Perspective With Directions for Research. In R. E. Stark (Ed.), *The Johnson /and] Johnson Baby Products Company pediatric Round Table: Vol. 4. Language behavior in infancy and early childhood: Proceedings* (pp. 127–239). New York: Elsevier/North-Holland.
- Newman, J. D. (2007). Neural circuits underlying crying and cry responding in mammals. *Behavioral and Brain Sciences*, 182(2), 155–165.
- Nishimura, T., Mikami, A., Suzuki, J., & Matsuzawa, T. (2003). Descent of the larynx in chimpanzee infants. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, 12(100), 6930–6933.
- Oller, D. K. (1986). Metaphonology and infant vocalizations. In B. Lindblom & R. Zetterström (Eds.), *Precursors of early speech: Wenner-Gren International Symposium Series. Proceedings of an International Symposium held at the Wenner-Gren Center, Stockholm* (Vol. 44, pp. 21–35).
- Oller, D. K. (2000). *The emergence of the speech capacity*. Mahwah, N.J: Lawrence Erlbaum Associates.
- Oller, D. K., Buder, E. H., Ramsdell, H. L., Warlaumont, A. S., Chorna, L., & Bakeman, R. (2013). Functional flexibility of infant vocalization and the emergence of language. *Proc. Natl. Acad. Sci. U.S.A.*, 110, 6318–6323.
- Oller, D. Kimbrough; Eilers, Rebecca E.; Basinger, Devorah; Steffens, Michele L.; Urbano, Richard (1995): Extreme poverty and the development of precursors to the speech capacity. In: *First Language* 15 (44), S. 167–187. DOI: 10.1177/014272379501504403
- Oller, D. K., Eilers, R., Neal, A., & Cobo-Lewis, A. (1998). Late onset canonical babbling: a possible early marker of abnormal development. *American Journal of Mental Retardation*, 103(3), 249–263.

- Oller, D. K., Eilers, R., Neal, A., & Schwartz, H. (1999). Precursors to speech in infancy: the prediction of speech and language disorders. *Journal of Communication Disorders*, 32(4), 223–245.
- Oller, D. K. (1980). The emergence of the sounds of speech in infancy. In G. K. J. Yeni-Komshian & A. Ferguson (Eds.), *Child phonology: Volume 1. Production* (pp. 93–112). New York: Acad. Pr.
- Pachtner, S. (2016). *Quantitative Charakterisierung akustischer und phonetischer Eigenschaften kanonischer Babblen von gesunden Kindern mit deutscher Umgebungssprache: Inaugural-Dissertation zur Erlangung der Doktorwürde.*
- Papousek, M.; Papousek, H. (1981): *Musical elements in the infant's vocalizations: their significance for communication, cognition and creativity.* In: L. P. Lipsitt (Hg.): *Advances in infancy research.* Norwood: Ablex (1), S. 163–224.
- Papousek, M. (1994). *Vom ersten Schrei zum ersten Wort: Anfänge der Sprachentwicklung in der vorsprachlichen Kommunikation.* Bern: Verlag Hans Huber.
- Papousek, M. (2014). Kommunikation und Sprachentwicklung im ersten Lebensjahr. In M. Cierpka (Ed.), *Frühe Kindheit 0-3 Jahre: Beratung und Psychotherapie für Eltern mit Säuglingen und Kleinkindern* (2nd ed., pp. 69–80). Heidelberg: Springer.
- Perani, D., Saccuman M. C., Scifo, P., Anwander, A., Spada, D., Baldoli, C.; Lohmann, G. & F. A. D. (2011). Neural language networks at birth. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, 108(38), 16056–16061.
- Ploog, D. (1979). Phonation, emotion, cognition, with reference to the brain mechanisms involved. *Ciba Foundation Symposium*, 69.
- Quast, A., Hesse, V., Hain, J., & Wermke, K. (2016). Baby babbling at five months linked to sex hormone levels in early infancy. *Infant Behavior & Development*, 44, 1–10.
- Riedl, R. (1975). *Die Ordnung des Lebendigen: Systembedingungen der Evolution.* Hamburg/Berlin: Parey.
- Robb, M. P., Chen, Y., & Gilbert, H. R. (1997). Developmental Aspects of Formant Frequency and Bandwidth in Infants and Toddlers. *Folia Phoniatrica*, 49, 88–95.

- Robb, M. P., Yavarzadeh, F., Schluter, P. J., Voit, V., Shehata-Dieler, W., & Wermke, K. (2020). Laryngeal constriction phenomena in infant vocalizations. *Journal of Speech, Language, and Hearing Research*, 63(1), 49-58.
- Rothgänger, H. (2003). Analysis of the sounds of the child in the first year of age and a comparison to the language. *Early Human Development*, 75(1-2), 55–69.
- Ruzza, B., Rocca, F., Lenti Boero, D., & Lenti, C. (2003). Investigating the Musical Qualities of Early Infant Sounds. *Annals New York Academy of Sciences*, 999, 527–529.
- Sachse, S. (2007). *Neuropsychologische und neurophysiologische Untersuchungen bei Late Talkers im Quer- und Längsschnitt: Dissertation zum Erwerb des Doktorgrades der Humanbiologie an der Medizinischen Fakultät der Ludwig-Maximilians-Universität zu München.*
- Sachse, Steffi; Bockmann, Ann-Katrin; Buschmann, Anke (2020): Sprachentwicklung. Berlin, Heidelberg: Springer Berlin Heidelberg, zuletzt geprüft am 09.06.2021
- Saint-Georges, C., Chetouani, M., Cassel, R., Apicella, F., Mahdhaoui, A., Muratori, F.; Cohen, D. (2013). Motherese in interaction: at the cross-road of emotion and Cognition. *Plos one*, 8(10), e78103.
- Sandner, G.W., & Wagner, E. (1981). Vocal Communication with a Three-Month Old Baby. *Child Language Development*. (20), 116–123.
- Sato, K., Hirano, M., & Nakashima, T. (2001). Fine Structure Of The Human Newborn And Infant Vocal Fold Muscosae. *Annals of Otology, Rhinology & Laryngology*, 110(5).
- Scheiner, E., Hammerschmidt, K., Jürgens, U., & Zwirner, P. (2002). Acoustic Analyses of Developmental Changes and Emotional Expression in Preverbal Vocalizations of Infants. *Journal of Voice*, 16(4), 509–529.
- Schweinfurth, J. M., & Thibeault, S. L. (2008). Does Hyaluronic Acid Distribution in the Larynx Relate to the Newborn´s Capacity for Crying? *The Laryngoscope*, 118, 1692–1699.

- Shehata-Dieler, W.; Ehrmann-Mueller, D.; Wermke, P.; Voit, V.; Cebulla, M.; Wermke, K. (2013): Pre-speech diagnosis in hearing-impaired infants: How auditory experience affects early vocal development. In: *Speech, Language and Hearing* 16 (2), S. 99–106
- Simonyan, K., & Horwitz, B. (2011). Laryngeal Motor Cortex and Control of Speech in Humans. *Neuroscientist*, 17(2), 197–208.
- Smith, B. L., & Oller, D. K. (1981). A Comparative Study of Pre-Meaningful Vocalizations Produced by Normal Developing and Down's Syndrome Infants. *Journal of Speech and Hearing Disorders*, 46, 46–51.
- Stark, R. E. (1978). Features of infant sounds: the emergence of cooing. *Journal of Child Language*, pp. 379–390.
- Stark, R. E. (1980). Stages of speech development in the first year of life. In G. K. J. Yeni-Komshian & A. Ferguson (Eds.), *Child phonology: Volume 1. Production* (pp. 73–90). New York: Acad. Pr.
- Stoel-Gammon, C. (1988). Prelinguistic Vocalizations of Hearing-Impaired and Normally Hearing Subjects: A Comparison of Consonantal Inventories. *Journal of Speech and Hearing Disorders*, 53, 302–315.
- Stoel-Gammon, C. (1989). Prespeech and early speech development of two late talkers. *First Language*, 9(6), 207–223.
- Stoel-Gammon, C. (1998). Role of Babbling and Phonology in Early Linguistic Development. In A. M. Wetherby, S. F. Warren, & J. Reichle (Eds.), *Transitions in prelinguistic communication* (pp. 87–110). Baltimore: Paul H Brookes Pub Co.
- Stoel-Gammon, C., & Otomo, K. (1986). Babbling Development of Hearing-Impaired and Normally Hearing Subjects. *Journal of Speech and Hearing Disorders*, 51(1), 33–41.
- Suchodoletz, W. (2011). Früherkennung von umschriebenen Sprachentwicklungsstörungen. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie*, 39, 377–385. Retrieved from <https://doi.org/10.1024/1422-4917/a000136>

- Thom, S., Hoit, J., Hixon, T., & Smith, A. (2006). Velopharyngeal Function During Vocalization in Infants. *Cleft Palate Craniofacial Journal*, 43(5), 539–546.
- Titze, I. R. (1994). *Principles of voice production*. Englewood Cliffs, N.J: Prentice Hall.
- Trevarthen, C. (1993). "The self-born in intersubjectivity: the psychology of infant communication". In U. Neisser (Ed.), *The Perceived Self: Ecological and Interpersonal Sources of Self-Knowledge* (pp. 183–229). New York: Cambridge University Press.
- Tucker, H. M. (1993). *The larynx* (2nd). New York, Stuttgart: Thieme Medical Publishers; Georg Thieme Verlag.
- Van de Ven, V., Esposito, F., & Christoffels, I. K. (2009). Neural network of speech monitoring overlaps with overt speech production and comprehension networks: a sequential spatial and temporal ICA study. *Neuroimage*, 47(4), 1982–1991.
- Velleman, S. L. (1994). The interaction of phonetics and phonology in developmental verbal dyspraxia, two case studies. *clinics in communication disorders*, 4(1), 66–77.
- Vihman, M. M. (1996). *Phonological development: The origins of language in the child*. Cambridge, Mass: Blackwell.
- Voit, V., Shehata-Dieler, W., Cebulla, M., & Wermke, K. *Acoustic Phenomena of Comfort Sounds – A Comparison of Healthy and Hearing Impaired Infants at 4 – 7 Months* (Poster presented at 14th ICPLA conference, Cork, Ireland, 2012).
- Vorperian, H. K., & Kent, R. D. (2007). Vowel acoustic space development in children: A synthesis of acoustic and anatomic data. *Journal of Speech, Language, and Hearing Research*, 50(6), 1510–1545. [https://doi.org/10.1044/1092-4388\(2007/104\)](https://doi.org/10.1044/1092-4388(2007/104))
- Vorperian, H. K., Kent, R. D., Gentry, L. R., & Yandell, B. S. (1999). Magnetic resonance imaging procedures to study the concurrent anatomic development of vocal tract structures: preliminary results. *International Journal Of Pediatric Otorhinolaryngology*. (49), 197–206.
- Vorperian, H. K., Kent, R. D., Lindstrom, M. J., Kalina, C.M., Gentry, L. R., & Yandell, B. S. (2005). Development of vocal tract length during early childhood: A magnetic resonance imaging study. *The Journal Of The Acoustical Society Of America*, 117(1), 338–350.

- Waltersbacher, A. (2020). *Heilmittelbericht 2020*. Berlin.
- Wasz-Höckert, O., Lind, J., Vuorenkoski, V., Partanen, T., & Valanne, E. (1968). *The infant cry, a spectrographic and auditory analysis*. London: Heinmann Medical Books.
- Weinrich, M., & Zehner, H. (2017). *Phonetische und phonologische Störungen bei Kindern: Aussprachetherapie in Bewegung* (5th ed.). Berlin: Springer.
- Wermke, K., Leising, D., & Stellzig Eisenhauer, A. (2007). Relation of melody complexity in infants' cries to language outcome in the second year of life: A longitudinal study. *Clinical Linguistics & Phonetics*, 21(11-12), 961–973. <https://doi.org/10.1080/02699200701659243>
- Wermke, K., & Mende, W. (2011). From emotion to notion: The importance of melody. In J. Decety & J. T. Cacioppo (Eds.), *The Oxford handbook of social neuroscience* (pp. 624–648). New York: Oxford University Press.
- Wermke, K. (2016). From Melodious Cries to Articulated Sounds: Melody at the Root of Language Acquisition. In M. C. Fonseca-Mora & M. Gant (Eds.), *Melodies, Rhythm and Cognition in Foreign Language Learning*. Cambridge Scholars Publishing.
- Wermke, K., Mende, W., Manfredi, C., & Brusciaglioni, P. (2002). Developmental aspects of infant's cry melody and formants. *Medical Engineering & Physics*, (24), 1119–1133.
- Wermke, K., Pachtner, S., Voit, V., Lamm, B., Kärtner, J., & Keller, H. (2013). Acoustic Properties of Comfort Sounds of 3-Month-old Cameroonian (Nso) and German Infants. *Speech, Language and Hearing*, 16(3), 149–162.
- Wermke, K., Mende, W. (1993). Variability of the cry melody as an indicator for certain developmental stages (Paper presented at the 15th Annual Conference of the IEEE Engineering in Medicine and Biology Society). San Diego California USA.
- Wermke, K., Mende, W. (1994). Ontogenetic development of infant cry- and non-cry vocalization as early stages of speech abilities. In R. Aulanko & A.M. Korpjaakko-Huuhka (Eds.), *Proceeding of the third congress of the international clinical phonetics and linguistics association* (pp. 181–189). Helsinki.

- Wermke, K. (2002). *Untersuchung der Melodieentwicklung im Säuglingsschrei von monozygoten Zwillingen in den ersten 5 Lebensmonaten*: Medizinische Fakultät - Universitätsklinikum Charité.
- Wermke, K., & Friederici, A. D. (2004). Developmental changes of infant cries--The evolution of complex vocalizations. *Behavioral and Brain Sciences*, 27(4), 474–475. <https://doi.org/10.1017/S0140525X04390102>
- Wermke, K. (2004). Vom Schreien zur Sprache - Was die Schrei-Melodien von Säuglingen über die vorsprachliche Entwicklung aussagen. *Frühförderung interdisziplinär*, 23(2).
- Wermke K., Linz C., Hasenberg A., Kunz F., Meyer-Marcotty P., Schweitzer T. (2017). Six months-old infants with deformational plagiocephaly do not differ from unaffected infants with respect to vocal control. *International Journal Of Pediatric Otorhinolaryngology*. (102), 15–20.
- Wirth, G. (2000). *Sprachstörungen, Sprechstörungen, Kindliche Hörstörungen: Lehrbuch für Ärzte, Logopäden und Sprachheilpädagogen* (5th ed.). Köln: Deutscher Ärzte-Verlag.
- Yang, C., Crain, S., Berwick, R., Chomsky, N., & Bolhuis, J. (2017). The growth of language: Universal Grammar, experience, and principles of computation. *Neuroscience and Biobehavioral Reviews*, 81, 103–119.
- Zheng, Z. Z., Munhall K. G., & Johnsrudel, I. S. (2010). Functional overlap between regions involved in speech perception and in monitoring one's own voice during speech production. *Journal of Cognitive Neuro Sciences*, 22(8), 1770–1781.
- Zollinger, B. (2015). *Die Entdeckung der Sprache* (9th ed.). Bern: Haupt.

Anhang

Tabelle 30: Eckdaten der Studien zu akustischen und phonetischen Eigenschaften in Komfortvokalisationen im Alter zwischen 3 und 7 Monaten

Autoren/ Jahr Umgebungssprache	Stichprobengröße (N) Alter der untersuchten Kinder Untersuchte Parameter	Vokalisationstyp- und Anzahl (n) Methode	Ergebnisse Bemerkungen
Bettany, 2004 englisch	N=1 1-6 Monate Phonation, Pitch, Duration	n=120 Auditiv-perzeptiv	Vokalisation: Äußerungen innerhalb einer Expirationsphase Kontinuierlicher Anstieg der mittleren Vokalisationslänge zwischen dem 1. Monat (0,44s) und 6. Monat (1,60s), Anstieg zwischen Monat 3 und 4 signifikant ($p < 0.05$)
Bloom, 1989 englisch	N= 80 3. Monat Lautlänge in Abhängigkeit vom sozialen Kontext	Komfortvokalisationen Auditiv-perzeptiv Unterteilung in Vokalisationen mit großer melodischer Variabilität (Silbencharakter, voll-resonant) und Vokalisationen mit gleichförmiger Melodie (eher nasal, Gurrlaute)	Vokalisationen mit Silbenstruktur traten seltener auf Sozialer Kontext hat keinen signifikanten Einfluss auf die Vokalisationslänge Laute mit Silbenstruktur sind länger (M= 1.22s) als vokalische Vokalisationen (M= 0.87s) Laute mit Silbenstruktur treten häufiger während verbalem Turn-Taking auf ABER Die Vokalisationslänge beider Vokalisationstypen zeigt keine

Autoren/ Jahr Umgebungssprache	Stichprobengröße (N) Alter der untersuchten Kinder Untersuchte Parameter	Vokalisationstyp- und Anzahl (n) Methode	Ergebnisse Bemerkungen
Clement, 2004 niederländisch	N=12 (6 normalhörend, 6 hörgeschädigt) 2-12 Monate Akustische Parameter, Silbenstruktur, Artikulationsort- und Art	Komfortvokalisationen n= 5381 längsschnittlich in Hinblick auf Gruppenunterschiede Spektralanalyse	signifikanten Veränderungen in Abhängigkeit vom sozialen Kontext Vokalisation innerhalb einer Expirationsphase; Mittlere F ₀ Gesamtzeitraum 294 Hz (SD 152 Hz); Signifikanter Alterseffekt durch Varianzanalyse; F ₀ max 430 Hz (SD 234) F ₀ min 294 Hz (SD 110) Mittlere Vokalisationslänge 915ms (kein Ausschluss von Cry), Anstieg mit steigendem Alter, signifikanter Peak bei 3,5 Monaten (1441ms)
D'Odorico, 1984 italienisch	N= 4 (2 männlich, 2 weiblich) 4-6 Monate Lautlänge in Abhängigkeit vom Melodiemuster und Vokalisationstyp	Längsschnitt Cry- and Non-Cry n= 744 Spektralanalyse	Definition Vokalisation: Pause > 50 ms! Non-Cry-Vokalisationen haben geringere F ₀ und Vokalisationslänge als Cry Vokalisationslänge insgesamt ist eher vom Melodietyp (rising, falling, level) als vom Vokalisationstyp abhängig (Cry, Non-Cry): flache Melodietypen sind immer kürzer als steigende oder fallende Melodietypen

Autoren/ Jahr Umgebungssprache	Stichprobengröße (N) Alter der untersuchten Kinder Untersuchte Parameter	Vokalisationstyp- und Anzahl (n) Methode	Ergebnisse Bemerkungen
D'Odorico et al., 1985 italienisch	N= 4 (2 männlich, 2 weiblich) 4-8 Monate Temporäre Eigenschaften	Cry and Non-Cry n=649 noncry (n= 93) Spektralanalyse Unterteilung der Vokalisationen in stimmhaft, halb-stimmhaft und stimmlos	Definition Vokalisation: Pause von mindestens 2s zwischen einzelnen Vokalisationen Länge stimmlose Comfort-Vok: 299- 623ms Länge stimmhafte Comfort-Vok: 518- 564ms Länge halb-stimmhafte Comfort-Vok: 666-798ms
D'Odorico & Franco, 1991 italienisch	N=5 4-11 Monate Akustische Eigenschaften in Relation zum sozialen Kontext	Komfortvokalisationen Spektralanalyse	Einzelne Vokalisationen sind durch Pausen von mindestens 300ms definiert
Delack, 1976 Delack & Fowlow, 1978 englisch	N=19 (7 weiblich, 12 männlich) 0-12 Monate Akustische Parameter	n= 11000 längsschnittlich	Geschlechtsspezifische Unterschiede in allen akustischen Parametern; Peaks der mittleren F ₀ mit 6 Monaten; Kaum Entwicklungsveränderungen; Mittlere F ₀ 355 Hz Anstieg der mittleren Vokalisationslänge um 50% im ersten LJ mit geschlechtsspezifischen Unterschieden und Peaks im 6. und 9. Monat
Hsu et al., 2000 englisch	N=13 0-6 Monate Sprechqualität (konsonantisch vs. vokalisches)	Komfortvokalisationen Auditiv-perzeptiv	Silben- und Melodiekomplexität nehmen parallel mit zunehmendem Alter zu

Autoren/ Jahr Umgebungssprache	Stichprobengröße (N) Alter der untersuchten Kinder Untersuchte Parameter	Vokalisationstyp- und Anzahl (n) Methode	Ergebnisse Bemerkungen
	Melodiekomplexität		
Iyer & Oller, 2008 englisch	N= 16 (8 normalhörend, 8 hochgradig hörgestört) 3 Zeitpunkte: 3 Mo vor kanonischer Phase, kanonische Phase, 3 Mo nach kanonischer Phase Suprasegmentale Eigenschaften	Komfortvokalisationen mit rein vokal-ähnlichem Kern und Vokalisationen mit konsonantischem Element n= 4342 (Silben-Basis) Gruppenvergleich auditiv-perzeptiv	Ergebnisse für normalhörende Kinder: Mittlere F ₀ = 299 Hz Kein Alterseffekt Kein Effekt des Vokalisationstyps
Kent & Murray, 1982 englisch	N= 21 3, 6, 9 Monate Mittlere F ₀ , Vokalisationslänge	Komfortvokalisationen Spektralanalyse (Berechnung durch Visi-Pitch bis auf 1 Hz genau) Querschnitt	Mittlere F ₀ variiert zwischen 445 Hz (3 Monate), 450 Hz (6 Monate) und 415 Hz (9 Monate); Interindividueller Variationsbereich 350-500 Hz keine genaue Definition des untersuchten Vokalisationstyps Hauptsächlich Kurzlaute (<0,4s), aber auch Vokalisationslängen zwischen 0,4 und 2s; Histogramme zeigen leichten Anstieg mit zunehmendem Alter
Laufer & Horii, 1977 englisch	N= 4 0-6 Monate Akustische Parameter	Komfortvokalisationen n= 787 Längsschnittlich	mittlere F ₀ (gesamt) von 335 Hz Variationsbereich 317-342 Hz Sehr geringe Veränderungen der mittleren F ₀ über den Untersuchungszeitraum

Autoren/ Jahr Umgebungssprache	Stichprobengröße (N) Alter der untersuchten Kinder Untersuchte Parameter	Vokalisationstyp- und Anzahl (n) Methode	Ergebnisse Bemerkungen
			Anstieg der mittleren Vokalisationslänge zwischen 2. Monat (600ms) und 6. Monat (1500ms)
Legerstee, 1991 englisch	N=8 0-6 Monate Melodiekomplexität in Bezug zu Mutter-Kind-Interaktion	Längsschnitt Auditiv-perzeptiv	Signifikant mehr melodische Strukturen mit steigendem Alter; Signifikant mehr melodische Strukturen, wenn aktive Mutter-Kind-Interaktion
Murry et al., 1983 englisch	N= 1 (weiblich) 0-4 Monate Suprasegmentale Eigenschaften in Relation zum Entwicklungsalter	Hunger-Schrei, Komfort-Schrei, Komfortvokalisation längsschnittlich	Definition der einzelnen Vokalisationen durch Pause von mindestens 0.5s; Erstmaliges Auftreten von Komfortvokalisationen im Alter von 8 Wochen; Mittlere F ₀ mit 8 Wochen 391.2 Hz Mittlere F ₀ mit 10 Wochen 294.1 Hz Mittlere F ₀ mit 12 Wochen 382.1 Hz Mittlere F ₀ Gesamtzeitraum 355.8 Hz Vokallänge bei Komfortvokalisationen nimmt kontinuierlich zu; Mittelwert 470ms
Rothgänger, 2003 deutsch	N=15 0-12 Monate Akustische Parameter (cry, babbling) im Vergleich zu Erwachsenensprache (male, female)	Schrei (n=583), Babbeln (n=201) längsschnittlich Spektralanalyse Babbling = „language-like sounds“ (keine genaueren Angaben zur Lautauswahl)	Mittlere F ₀ für Babbel-Vokalisationen sinkt im Untersuchungszeitraum signifikant von 389.3 Hz auf 336.9 Hz ab Mittlere F ₀ Gesamtzeitraum 342.7 Hz, SD 51 Hz

Autoren/ Jahr Umgebungssprache	Stichprobengröße (N) Alter der untersuchten Kinder Untersuchte Parameter	Vokalisationstyp- und Anzahl (n) Methode	Ergebnisse Bemerkungen
			Mittlere Vokalisationslänge 553,1ms, SD 341,4 Kontinuierliche Zunahme der Vokalisationslänge zwischen 1. Monat (406,4ms) und 12. Monat (699,1ms)
Ruzza et al., 2003 marokkanisch italienisch	N= 5 (3 marokkanisch, 2 italienisch) 2. Lebensmonat Mittlere F ₀ , Melodiekontur, Vokalisationslänge	Prothone (Gurren, Babbeln) n= 109 marokkanisch n= 54 italienisch Gruppenvergleich	Mittlere F ₀ (gesamt) von 364, 35 Hz Keine Gruppenunterschiede bezüglich der untersuchten Parameter Mittlere Vokalisationslänge: 730ms
Scheiner et al., 2002 deutsch	N= 7 Erstes Lebensjahr Suprasegmentale Eigenschaften in Relation zum emotionalen Status	Cry, Babbeln n= 16.322	Keine signifikanten altersbedingten Veränderungen der mittleren F ₀ Duration 727,4 ms (marokkanisch) und 740,1 ms (italienisch)
Voit, Shehata-Dieler, Cebulla, & Wermke, 2012 deutsch	N=12 (6 normalhörend, 6 hörgeschädigt) 4-7 Monate Akustische Parameter	Komfortvokalisationen n= 461 Spektralanalyse	Mittlere F ₀ 293 Hz
Wermke et al., 2013 deutsch Nso (Kamerun)	N= 33 (14 deutsche, 19 nso) 3. Lebensmonat Akustische Eigenschaften	Komfortvokalisationen n= 808 Spektralanalyseverfahren	Mittlere F ₀ (deutsch) 318 Hz (SD 72) Mittlere F ₀ (nso) 355 Hz (SD 98 Hz) Kein signifikanter Gruppenunterschied; stützt Ergebnisse von Ruzza et al. (2003) Mittlere Vokalisationslänge 999ms

Tabelle 31: Ergebnisse der standardisierten Testverfahren zur sprachlichen Entwicklung. (ua: unauffällig; a: auffällig, rot unterlegt).

Kind	VB	VD	VE	VF	VG	VH	VI	VL	AC	AD	AK
ELFRA-1											
12 Monate Sprachproduktion	ua	ua	ua	a	ua	a	ua	ua			
Sprachverständnis	ua	ua	ua	ua	ua	ua	ua	ua			
Gesten	ua	ua	ua	ua	ua	ua	ua	ua			
Mimik	ua	ua	ua	ua	ua	ua	ua	ua			
ELFRA-1											
18 Monate Sprachproduktion	ua		ua	ua	ua	ua	ua	ua			
Sprachverständnis	ua		ua	ua	ua	ua	ua	ua			
Gesten	ua		ua	ua	ua	ua	ua	ua			
Mimik	ua		ua	ua	ua	ua	ua	ua			
SETK-2											
Alter (Monate)	31		34	32		32			29		29
T-Werte											
Verstehen I	69		62	62		62			54		45
Verstehen II	54		65	65		65			65		49
Produktion I	66		71	56		61			73		47
Produktion II	49		51	38		49			54		41
SETK-3-5											ua

Danksagung

Nachfolgend möchte ich mich bei denjenigen Personen bedanken, die durch ihre Mithilfe ganz entscheidend zum Gelingen meiner Doktorarbeit beigetragen haben. Ohne sie wäre ein Abschluss meiner Dissertation nicht zustande gekommen.

Zunächst möchte ich mich bei meinem Erstgutachter Prof. Dr. Detlef Hansen für die Unterstützung meines Dissertationsprojekts und seine konstruktive Kritik in diesem Zusammenhang bedanken.

Mein herausragender Dank gilt meiner Zweitgutachterin Frau Prof. Dr. Kathleen Wermke. Durch ihre jahrelange wissenschaftliche Begleitung meiner Arbeit, ihre Gedanken und gemeinsame fachliche Diskussionen mit ihr, ist diese Arbeit erst möglich geworden. Ihr unermüdlicher Forschungsdrang und ihre Begeisterung für die Wissenschaft haben mich angesteckt und somit ganz wesentlich zu meinem Verständnis wissenschaftlicher Zusammenhänge beigetragen. Sie hat mich durch ihre lieben Worte stets motiviert weiterzumachen und ist mir über die Jahre hinweg nicht nur in fachlichen Dingen mit Rat und Tat zur Seite gestanden.

Bei Peter Wermke möchte ich mich für die großzügige Unterstützung bei der Datenaufbereitung und Datenarchivierung bedanken. Ohne seine fundierten fachlichen Kenntnisse, sein Ideenreichtum und seinen großen Zeiteinsatz wäre die Analyse der enormen Datenmengen nicht möglich gewesen.

Mein weiterer Dank gilt den Familien und Kindern, durch deren kontinuierliche Anfertigung von Lautaufnahmen und ihrem Einverständnis der Verwendung der Daten für wissenschaftliche Zwecke das Projekt überhaupt erst zustande kam.

Meinen ehemaligen Kolleginnen Dr. Birgit Mampe-Keller, Dr. Annette Prochnow, Dr. Daniela Eiband und Dr. Sandra Michel danke ich für die unvergessliche gemeinsame Promotionszeit am ZVES. Auch bei meinem Team am Lehrstuhl für Sprachheilpädagogik möchte ich mich für die motivierende und kollegiale Unterstützung und Beratung bedanken.

Nicht zuletzt bedanke ich mich von Herzen bei meiner Familie und meinen engsten Freunden. Ohne ihr Verständnis, ihre Aufmunterungen und die liebevolle Betreuung meiner Kinder in intensiven Arbeitsphasen wäre das Projekt nicht möglich gewesen.