

Aus dem Zentrum für Zahn-, Mund- und Kiefergesundheit
des Universitätsklinikums Würzburg
Klinik für Kieferorthopädie
Direktorin: Professor Dr. med. dent. A. Stellzig-Eisenhauer

**Melodisch-prosodische Eigenschaften von Vokanten
bei normalhörenden und hochgradig sensorineural hörgestörten Säuglingen
im Alter von 3-6 Monaten**

Inauguraldissertation

zur Erlangung der Doktorwürde
der Medizinischen Fakultät
der
Julius-Maximilians-Universität Würzburg

vorgelegt von
Fabian Clad
aus Freiburg im Breisgau

Würzburg, 23.12.2022

Referentin: **Univ.-Prof. Dr. rer. nat. Kathleen Wermke**
Koreferentin: **apl. Prof. i.R. Dr. med. Wafaa Shehata-Dieler**
Berichterstattung: **Univ.-Prof. Dr. rer. nat. Grit Hein**
Dekan: **Prof. Dr. Matthias Frosch**

Tag der mündlichen Prüfung: 03.11.2023

Der Promovend ist Zahnarzt

Inhaltsverzeichnis

Inhaltsverzeichnis.....	I
1 Einleitung	1
2 Zielsetzung	8
3 Material und Methoden	9
3.1 Ethikvotum.....	9
3.2 Design der Kohortenstudie	9
3.3 Einbettung der vorliegenden Arbeit in das Gesamtprojekt.....	11
3.4 Untersuchungskollektiv	13
3.4.1 Patientengruppe (HI)	13
3.4.2 Kontrollgruppe	14
3.5 Übersicht der Aufnahmetage und der Anzahl aufgezeichneter Laute	15
3.6 Aufbereitung und Analyse der Komfortvokalisationen	16
3.6.1 Von der Audiodatei zur Langserie (Datenaufbereitung).....	17
3.6.2 Von der Langserie zum Einzelsignal: Schnitt der Langserien und Segmentierung der Einzelereignisse	17
3.6.3 Extraktion der egressiven Komfortvokalisationen („e“) und Übergabe der Audiodateien und f0-Werte an CDAP	18
3.7 Qualitative Analyse und audiovisuelle Identifizierung der Vokanten.....	20
3.7.1 Interaktive f0-Kontrolle	20
3.7.2 Bestimmung des Vokalisationstyps	21
3.8 Validierung der Vokanten	24
3.8.1 Breitbandspektogramm-Analyse	24
3.8.2 Formantanalyse	27
3.9 Melodiestrukturanalyse einfacher Vokanten	30
3.9.1 1B und Kurzlaute (einfache Struktur)	30
3.9.2 Doppelbögen (2B) und segmentierte Strukturen (SEG) (Komplexe Strukturen).....	31
3.10 Vermessung der prosodie-relevanten Melodieeigenschaften	32
3.11 Statistik.....	33
4 Ergebnisse	38
4.1 Formantanalyse	38
4.2 Melodiestrukturanalyse	42
4.3 Melodisch-prosodische Eigenschaften von Vokanten	45
4.3.1 Melodielängen aller Vokanten.....	45
4.3.2 Melodielänge in einfach strukturierten Vokanten (1B).....	50

4.3.3	Melodielänge der Vokanten mit doppelböigen Melodien (2B)	54
4.4	Rhythmusanalyse in doppelböigen Vokanten	56
4.4.1	Längen-Verhältnisse LB1 zu LB2 und Rhythmik-Charakteristika	56
4.4.2	Bogenform-Kombinationen in doppelböigen Vokanten.....	57
5	Diskussion	58
5.1	Anteil der Vokanten im Lautrepertoire.....	58
5.2	Formantanalyse	60
5.3	Melodiestrukturanalyse	66
5.4	Dauer von Vokanten	70
5.5	Melodisch-Rhythmische Eigenschaften in doppelböigen Vokanten	74
5.6	Formkombinationen in doppelböigen Vokanten	78
6	Limitationen	80
7	Zusammenfassung	81
8	Literaturverzeichnis	83
9	Eigene Publikationen	93
10	Appendix	
I.	Abkürzungsverzeichnis.....	
II.	Abbildungsverzeichnis	
III.	Tabellenverzeichnis.....	
IV.	Danksagung	
V.	Lebenslauf	

1 Einleitung

Die linguistische Fachliteratur beschreibt verschiedene Stadien von sogenannten Komfortvokalisationen in der vokalen Entwicklung im ersten Lebensjahr. Diese beinhalten die Phonationsphase, die Gurr-Phase (Cooing), das Babbeln (Babbling) und das sogenannte Jargon-Babbling (Papoušek 2001). Die Phasen sind durch Säuglingsvokalisationen mit typischen Klangeigenschaften charakterisiert, die den Phasen auch ihren Namen gegeben haben. Während vor allem die Forschergruppe um Kimbrough Oller (USA) ein Verfechter dieses Phasenmodells für die vorsprachliche Analyse ist (Oller 1980; Oller 2000), weisen andere Autoren darauf hin, dass diese Phaseneinteilung relativ unpräzise und insbesondere für klinische Fragestellungen ungeeignet ist. Es gibt Unstimmigkeiten bezüglich der phasenspezifischen Charakteristiken, dem Auftretsalter oder der Dauer der einzelnen Phasen, insbesondere der Babbelphase (Bauer 1988; Kent und Bauer 1985; Papoušek 2001; Smith et al. 1988). Bauer (1988) fasst die Nachteile des Phasenmodells zusammen:

„Characterizing phonetic development with such a stage model is problematic. Kent (1983) reviewed the stage model and highlighted its rarely supported assumptions. The requirements of unitary and well-defined stages, as well as their orderly occurrence, are rarely met (e.g. Smith, B., 1988). Imposing definite categories upon the continuous processes of phonetic development is further complicated by individual differences in the expression of ‘stages’. Its related requirement of equilibration, or the notion that unitary plateaus of developmental stability exist in a ‘stage’, would be particularly difficult in application to variable babbling vocalizations.” (Bauer 1988, S. 350)

Obwohl eine sequenzielle Abfolge der Organisationsstadien der vokalen Entwicklung hilfreiche Etikettierungen für den Klangeindruck liefern, ist diese eine sehr starke Vereinfachung. Vorsprachliche Analysen, basierend auf einer zu starken Fokussierung auf das Erreichen bzw. Nichterreichen von Stadien, können verursachen, dass der Beobachter eine Vielzahl bedeutsamer, aber vielleicht nicht ins strenge Schema

passende, Lautelemente unberücksichtigt lässt und die Entwicklung nicht adäquat erfasst. In der klassischen Literatur zum Stadienmodell werden z.B. Gurr-Laute als „einfache Komfortlaute“ im Alter von 6 bis 16 Wochen beschrieben (Stark 1980). Schaut man sich diesen Altersbereich aber genauer an, dann gibt es eine Vielzahl von vokalen Ereignissen, die die Laute der Säuglinge in diesem Alter charakterisieren, die man nicht als „einfach“ beschreiben kann. Einfach sind allein die Vokanten in diesem Alter. Aber es gibt andere Lautproduktionen hoch komplexer Natur in diesem Alter. Ähnlich verhält es sich mit dem sogenannten marginalen und kanonischen Babbeln, das sehr variabel in seinem Auftretsalter ist und ein Überlappen dieser Phasen eher typisch als differenzierend ist. Insbesondere bei Risikokindern für die Sprachentwicklung kann man den Beginn der Babbelphasen aufgrund einer hohen inter-individuellen Variabilität nicht zuverlässig als Kriterium im Rahmen diagnostischer Fragestellungen verwenden. Dies wird aber bis heute in Studien dennoch getan, insbesondere wenn die vorsprachliche Entwicklung hörgeschädigter Säuglinge eingeschätzt und beschrieben wird (Eilers und Oller 1994; Iyer und Oller 2008; Oller et al. 1999).

Das Erreichen bestimmter vokaler „Phasen“ als Diagnosekriterium wird in der Regel auch deshalb oft unkritisch übernommen, weil dies durch die infraphonologische Theorie von Oller scheinbar validiert wurde (Buder et al. 2013; Oller 2000). Dabei wird übersehen, dass die Arbeitsgruppe um Oller in den vergangenen Jahren wiederholt Modifikationen an ihren Theorien vornehmen musste, weil sie kritischen Prüfungen nicht standhielt (Oller 2000). Der Hauptgrund dafür, dass für diagnostische Zwecke eine alleinige Phasenzuordnung frühkindlicher Äußerungen nicht ausreicht, ist die enorm große inter-individuelle Variabilität, die auf genetischen und exogenen Faktoren basiert. Die vokale Entwicklung jedes Säuglings verläuft zwar nach universalen Regeln, aber die Entfaltung der angeborenen Entwicklungsgesetzmäßigkeiten ist eng an individuelle medizinisch-biologische und sozio-emotionale Konditionen geknüpft. Zudem ist das auditiv-vokale System ein in kontinuierlicher Reifung befindliches System, das auf exogene Einflüsse besonders sensibel reagiert.

Basierend auf diesem theoretischen Hintergrund wurden die in der vorliegenden Arbeit analysierten Lautäußerungen von Säuglingen nicht einer bestimmten „Phase“ zugeordnet, sondern phasenübergreifend, anhand ihrer akustischen Merkmale analysiert. Es handelt sich um vokalähnliche Äußerungen für die in der Fachliteratur auch oft der Begriff „Vokant“ (engl. vocant) verwendet wird (Kent und Murray 1982; Martin 1981). Sie treten vor allem in den ersten Lebensmonaten auf, sind aber phasenübergreifend auch später noch zu beobachten (Buder et al. 2013).

Die Vokanten sind neben dem natürlichen Weinen typische Säuglingslaute, die noch keine supralaryngealen konsonantischen Vorstufen enthalten. Die konsonantischen Vorstufen bezeichnet man analog zu den Vokanten als „closants“ (Martin 1981). Diese vokalähnlichen Lautäußerungen sind also zunächst Ergebnis der reinen Phonation, erzeugt ohne intentionale artikulatorische Aktivität. Andere, im Untersuchungsalter erzeugte Äußerungen können bereits konsonantenähnliche, artikulatorische Elemente enthalten, also Kombinationen aus „vocants“ und „closants“ darstellen (Bauer 1988).

In der vorliegenden Arbeit wurden nur Vokanten untersucht. Sie wurden gewählt, weil sie die einzige Lautkategorie bilden, die kaum einem Entwicklungseffekt unterliegt. Sich im Untersuchungsalter noch in Reifung befindliche, artikulatorische Prozesse können die akustischen Merkmale der Vokanten nicht beeinflussen. Damit werden individuelle Reifeunterschiede der Vokaltraktentwicklung als Einflussfaktor ausgeschlossen und der Einfluss der auditorischen Rückkopplung bzw. ihres Fehlens bei der Lautproduktion hörgeschädigter Säuglinge besser separierbar. Auch wenn es im Verlauf der vokalen Entwicklung bei anderen Lautkategorien (Vokalisationstypen) zum Einbau von Konsonanten in vokalähnliche Strukturen kommt, wenn Registerwechsel trainiert werden oder mit verschiedenen Artikulationszonen exploriert wird, bleiben Vokanten definitionsgemäß erzeugt mit neutraler Vokaltraktposition, davon unbeeinflusst. Das macht sie zu einem geeigneten Untersuchungsobjekt in einem Lebensalter, das ansonsten durch ständige Veränderungen, bedingt durch Wachstum- und Reifeprozesse, charakterisiert ist. Sollten sich signifikante Unterschiede in bestimmten Eigenschaften zwischen normalhörenden und hörgestörten Säuglingen zeigen, wären

Vokanten auch klinisch einfach und reproduzierbar zu analysieren und könnten als Kandidat für die Frühdiagnose von Auffälligkeiten der frühen vokalen Entwicklung weiter untersucht werden.

Allerdings ist es nicht ganz so einfach wie dargestellt, da die Untersuchung von Vokanten auch laryngeale, prä-artikulatorische Mechanismen berücksichtigen muss. Es gibt inzwischen in der Fachliteratur eine veränderte Perspektive auf den Beginn der artikulatorischen Entwicklung. Im Unterschied zu früheren Annahmen, dass die frühe Artikulation erst mit dem Absinken des Kehlkopfes beginnt (Sasaki et al. 1977), geht man inzwischen davon aus, dass bereits von Anfang an zumindest prä-artikulatorische Prozesse in der Lautproduktion des Säuglings eine Rolle spielen. Bereits im natürlichen Weinen von Säuglingen gibt es Rhythmisierungen der Melodie (Grundfrequenzkontur) durch vollständige oder unvollständige Unterbrechungen des Luftstroms im Larynx (Wermke und Mende 2009).

Diese laryngealen Konstriktionen werden durch Unterbrechung der Stimmlippenschwingungen und durch Bewegungen aryepiglottischer Falten erzeugt (Esling 2005; Schenk 2022). Da sie als Vorstufe der späteren Silbenrhythmik angesehen werden, beginnt die artikulatorische Entwicklung des Säuglings also im Larynx (Robb et al. 2020). Dieses neue Modell des laryngealen Artikulators wurde von dem kanadischen Linguisten John Esling eingeführt und als „Laryngeal Articulator Model“ (LAM)- Modell bezeichnet (Esling 2005).

Das LAM-Modell modifiziert die vereinfachte Ansicht, dass sich die Lautproduktion in Phonation und Artikulation unterscheiden lässt. Diese klassische Unterscheidung basiert auf dem Quelle-Filter-Modell von Gunnar Fant (Fant G. 1960). In seinem Modell betrachtet Fant die an den Stimmlippen erzeugten Schwingungen als Ausgang einer punktförmigen Schallquelle (= Phonation). Der unmittelbar nachgeschaltete Vokaltrakt wird als lineares Sprachrohr verstanden. Durch Reflektion entlang dieses linearen Modells entstehen Resonanzen und Knotenpunkte, wodurch das Ausgangssignal verstärkt bzw. gefiltert wird (= Artikulation).

Man weiß inzwischen, dass dieses Modell die Verhältnisse der menschlichen Lautproduktion zu vereinfacht beschreibt. In verschiedenen Untersuchungen wurde gezeigt, dass sich unterschiedlichste nicht-lineare Interaktionen zwischen der Quelle und dem Filter beobachten lassen (Titze 1994). Diese Interaktionen gibt es, so nimmt man an, bereits von Anfang an, also auch in den frühesten Lautproduktionen der Säuglinge (Mende et al. 1990). Das LAM-Modell von Esling versucht diese Nichtlinearitäten zu berücksichtigen, beschreibt aber im Wesentlichen zwei artikulatorische Untersysteme, nämlich das laryngeale und das supralaryngeale bzw. orale (Esling 2005; Esling et al. 2019). Entgegen Fants Modell finden sich nach Esling bereits an den Stimmlippen Artikulationsphänomene in Form von laryngealen Konstriktionen (engl. *glottal stop*) (Esling 2005; Yavarzadeh 2020).

Untersucht man Säuglingslaute basierend auf dem LAM-Modell beginnt die artikulatorische Entwicklung früher als bisher angenommen, faktisch gleichzeitig mit der phonatorischen Entwicklung. Benner und Esling (2019) haben Säuglinge im Alter ab dem dritten Lebensmonat aus verschiedenen Sprachkulturen untersucht und laryngo-artikulatorische Lautelemente in Form verschiedener Konstriktionstypen beschrieben (Esling et al. 2019). Robb et al. (2020) und Yavarzadeh (2020) haben noch früher, bei Säuglingen in den ersten drei Lebensmonaten im natürlichen, spontanen Weinen ähnliche Konstriktionen beobachtet und als regelhaft auftretend beschrieben. Daher wurde die Definition der frühesten Komfortlaute, der Vokanten, erweitert und die laryngealen Konstriktionen einbezogen. Unter Vokanten, also vokalähnlichen Lautäußerungen wurde in der vorliegenden Arbeit daher alle egressiven Äußerungen verstanden, die mit neutraler Vokaltraktposition erzeugt wurden, aber laryngeale Konstriktionen enthalten können.

In der vorliegenden Arbeit wurden damit Vokanten mit und ohne laryngeale Konstriktionen untersucht und dabei analysiert, ob diese frühen Komfortvokalisationen bereits Unterschiede im Vergleich zwischen normalhörenden und hochgradig sensorineural hörgeschädigten Säuglingen aufweisen. Wie diese Vokanten in der

vorliegenden Arbeit, identifiziert und durch objektive Analysen verifiziert wurden, ist in den Kapiteln 3.7 & 3.8 ausführlich erklärt.

Im Rahmen eines Kooperationsprojektes wurden parallel zur pädaudiologischen Diagnostik die Lautaufnahmen am Comprehensive Hearing Center (CHC) der Klinik und Poliklinik für Hals-, Nasen- und Ohrenkrankheiten, plastische und ästhetische Operationen des Universitätsklinikums Würzburg unter Leitung von Frau Professorin Shehata-Dieler durchgeführt und anschließend im Zentrum für vorsprachliche Entwicklung und Entwicklungsstörung (ZVES) unter Leitung von Frau Professorin Wermke analysiert und ausgewertet.

Laut Studienprotokoll handelt es sich um eine Kohortenstudie, die mit einmaligem pädaudiologischem Untersuchungstermin mittels akustisch evozierter Potenziale (brainstem evoked response audiometry (BERA)) und einmaliger Komfortlautaufzeichnung zum Zeitpunkt des Neugeborenen-Hörscreenings 2 (NHS-2) bei fehlendem BERA-Erfolg im Neugeborenenhörscreenings 1 (NHS-1) konzipiert wurde. Alle Säuglinge, die im Untersuchungszeitraum der Kohortenstudie im NHS-1 auffällig waren, wurden in die Kohorte eingeschlossen. Das spezifische Ziel der Kohortenstudie ist die Untersuchung der vorsprachlichen (Komfort-) Lautproduktionen unter Berücksichtigung der Hörfähigkeiten zum Zeitpunkt des NHS-2. Danach wurden Säuglinge mit verschiedenen Graden ein- oder beidseitiger sensorineuraler Hörstörung, Säuglinge mit Schalleitungsstörung und normalhörende Säuglinge unterschieden. In der vorliegenden Arbeit, die zu Beginn der Auswertungen erstellt wurde, erfolgte der Vergleich normalhörender Säuglinge zu hochgradig sensorineural hörgeschädigten Säuglingen. Dieser Vergleich wurde gewählt, um Einflussfaktoren durch individuelle Hörleistungen zu minimieren.

Das Neugeborenenhörscreening basiert auf einem objektiven Messverfahren, der Messung auditiv evozierter Hirnstammpotenziale (AEHP, engl. BERA) und wird routinemäßig durchgeführt. Hierbei wird ein weiches Mikrofon in das Ohr des Neugeborenen eingeführt und über dieses Klicksignale gesendet (Cebulla et al. 2014; Shehata-Dieler et al. 2013). Mittels angebrachter Elektroden an der Kopfhaut werden

zeitgleich die akustisch evozierten Hirnstammpotenziale aufgezeichnet und somit die Funktionsfähigkeit des Innenohres und der nachgeschalteten Hörbahn im Hirnstamm ermittelt (Cebulla und Shehata-Dieler 2012). Die Durchführung ist non-invasiv und birgt keinerlei Risiken für das Neugeborene. Ein Vorteil ist zudem, dass die Messungen keine aktive Mitarbeit erfordern und uneingeschränkt und problemlos auch am schlafenden Neugeborenen durchgeführt werden können. Das am UKW verwendete Messverfahren mit Hilfe des BERAPHON© weist eine nahezu 100%ige Sensitivität, wie auch gegenüber anderen Messverfahren erhöhte Spezifität von 97% auf (Cebulla et al. 2014). Die Kohortenstudie wurde konzipiert, um zukünftig im Screening nicht allein die Hörfunktion, sondern auch die Lautproduktion einzubeziehen.

Das ist klinisch insofern relevant, als dass im NHS-1 nicht ausschließlich Neugeborene mit einer sensorineuralen Hörbehinderung im Screening auffällig werden. Faktoren, wie Flüssigkeit im Gehörgang, der affektive Zustand beim Messen oder einer Hörkanal-Blockade durch einen harmlosen Propf, das Gestationsalter, der Geburtsmodus und das Geschlecht können zu einer Fehlmessung (falsch-positive Fälle) beim NHS-1 führen (Da Silva et al. 2015; Nakamura et al. 1985; Schwarz et al. 2017; Shapiro und Nakamura 2001; Stipdonk et al. 2016; Stuart 2020). Ein auffälliges NHS-1 ist demnach nicht unmittelbar gleichbedeutend mit einer persistierenden Hörstörung (Cebulla et al. 2014). Daher muss routinemäßig bei NHS-1 auffälligen Säuglingen ein weiteres Hörscreening (NHS-2) im Alter von 2-3 Monaten erfolgen (Abbildung 1).

Könnte man die Lautproduktionsleistungen zukünftig beim Neugeborenen-Hörscreening einbeziehen, wäre neben einer Senkung der Rate falsch-positiver Säuglinge auch eine zuverlässigere frühe Identifikation angeborener Hörstörungen zu erwarten. Die Kohortenstudie, in deren Rahmen die vorliegende Arbeit angefertigt wurde, dient diesem Ziel.

2 Zielsetzung

Ziel der vorliegenden Arbeit war die Untersuchung der melodisch-prosodischen Eigenschaften von Vokanten innerhalb der Komfortlautbildungen. Es galt einen Vergleich zwischen normalhörenden und hochgradig sensorineural hörgestörten Säuglingen vorzunehmen und dabei die Frage zu beantworten, inwieweit das Hörvermögen Einfluss auf Eigenschaften der primitivsten Form der Komfortlaute hat. Es geht um eine erste Hypothesengenerierung in einem bislang völlig unerforschten Untersuchungsbereich.

In der Literatur wird für ältere Kinder und Erwachsene angegeben, dass sich bestimmte prosodische Eigenschaften der Sprache hochgradig hörgestörter Patienten und Patientinnen gegenüber normalhörenden unterscheiden (Chesters et al. 2015; Iversen et al. 2008; Robb und Pang-Ching 1992; Trainor und Adams 2000). Dabei zeigt sich unter anderem, dass die hörgeschädigten Probanden und Probandinnen längere Wort- und Silbenlängen beim Sprechen verwenden. Hier sollte untersucht werden, ob sich temporale Unterschiede bereits in den Vokalisationslängen der primitivsten vorsprachlichen Vokalisationsform – Vokanten – finden lassen. Dies sollte unter Berücksichtigung der spezifischen Melodiestructur analysiert werden. Dazu wurden die zeitlichen Verläufe der Grundfrequenz (Melodie) nach der Anzahl enthaltener Bögen und Pausen differenziert (Kap. 3.9). Die Melodiestructur wurde nach ein-, zwei- und mehrböigen Formen kategorisiert.

Diese Forschungsarbeit bedurfte neben einer Formantanalyse zur Einordnung der identifizierten Vokanten im Vokalraumdreieck einer objektiven Lautsignalanalyse mit Hilfe geeigneter Programme. Diese standen im Labor am ZVES zur Verfügung. Folgende Fragstellungen in Bezug auf die rhythmisch-prosodischen Melodieeigenschaften von Vokanten sollten im Rahmen der Zielsetzung bearbeitet werden:

- (1) Welche Melodiestructuren zeigen Vokanten und mit welcher Häufigkeit treten diese im Untersuchungsalter von 2-6 Monaten auf?
- (2) Wie sehen die Verteilungsmuster der Melodielängen in Vokanten aus?
- (3) Welche rhythmischen Muster zeigen doppelböige Vokanten?
- (4) Unterscheiden sich die in (1) – (3) genannten Messgrößen zwischen den beiden Untersuchungsgruppen? Inwieweit beeinflusst das Hörvermögen die Vokanteneigenschaften?

3 Material und Methoden

3.1 Ethikvotum

Die vorliegende Arbeit war in ein Projekt mit dem Titel „Analyse von Komfortlauten (babbling) bei Säuglingen zum Zeitpunkt des Neugeborenen-Hörscreening-2“ (Beginn Januar 2018) eingebettet. Die Studie wurde unter Einhaltung der Deklaration von Helsinki durchgeführt (Ethikvotum-Nummer: 308/17 me). Es handelt sich um eine Kohortenstudie mit dem Ziel der Untersuchung sogenannter Komfortlaute (Nichtschrei-Vokalisationen) in der vorsprachlichen Entwicklung bei normalhörenden Säuglingen und Säuglingen mit sensorineuralen Hörstörungen. Dazu wurden in Zusammenarbeit mit einem Team von insgesamt sieben Doktoranden und Doktorandinnen Lautaufnahmen durchgeführt und anschließend labortechnisch analysiert. Der Autor war von Beginn an aktiv an der Studie beteiligt. In der vorliegenden Arbeit werden Teilergebnisse der Kohortenstudie berichtet.

3.2 Design der Kohortenstudie

Die Eltern der Säuglinge wurden über die Art und den Zweck der Studie mündlich aufgeklärt und, sofern die Einwilligung zur Teilnahme gegeben wurde, anhand zweier Informations- und Fragebögen während der Wartezeit spezifisch aufgeklärt. Die Teilnahme wurde in Form einer schriftlichen Einwilligungserklärung gegeben. Da, neben der Nicht-Einwilligung zur Teilnahme, keine weiteren Ausschlusskriterien galten, lag ein sehr heterogenes Probandenkollektiv mit divergierenden Begleiterkrankungen und familiären genetischen Dispositionen in der Kohorte vor.

Die Auswertung erfolgt daher zunächst in homogenen Untergruppen der Kohorte im Rahmen von Dissertationen der beteiligten Doktorandinnen und Doktoranden.

Für die vorliegende Arbeit wurde aus den ersten konsekutiven Aufnahmen eine Kontrollgruppe normalhörender Säuglinge ausgewählt. Dabei handelt es sich um eine Gruppe der falsch-positiven NHS-1-Fälle (Abbildung 1). Diese sollten mit einer Gruppe von Säuglingen verglichen werden, die eine hochgradige sensorineurale Hörminderung zeigten.

Der Autor war an den Lautaufnahmen nicht beteiligt, hat aber zu Projektanfang sämtliche aufgezeichneten Laute, über die hier in der Arbeit analysierten hinaus, wie in den folgenden Kapiteln 3.6 bis 3.10 beschrieben, vorverarbeitet und archiviert. Nach Vorliegen einer Minimaldatenmenge für die eigene Arbeit hat er mit der speziellen Auswertung des sogenannten Vokanten begonnen (vgl. 3.7). Zur Vollständigkeit und zum besseren Verständnis des Projektes wird die Erhebung der Lautaufnahmen bei den Säuglingen hier kurz erklärt.

Für die Lautaufnahmen erhielten die beteiligten Doktorandinnen eine Einweisung in die Handhabung des Aufnahmeegeräts (TASCAM, Modell DR-100) sowie Instruktionen für einen standardisierten Aufnahmevorgang unter Berücksichtigung der klinischen Gegebenheiten und Rücksichtnahme auf Eltern und Kind. Es sollte ein konstanter Abstand des Mikrofons zum Säugling eingehalten werden sowie alle potenziellen Störfaktoren und Hintergrundgeräusche auf ein Minimum reduziert werden.

Die Aufgabe während der Aufnahmen bestand im Wesentlichen darin, den Säugling in „Lautierlaune“ – durch eine „face-to-face“-Interaktion zu versetzen und zu entsprechenden Komfortlautäußerungen zu animieren (Hsu et al. 2001). Der Säugling sollte dafür in Rückenlage mit Blickkontakt zu einem der Kontaktpersonen (Mutter, Vater) befinden. Als Hilfsmittel konnten hierzu Spielobjekte (Attrappen) hinzugezogen werden. Da die spontane Lautproduktion der Säuglinge nur bedingt beeinflussbar ist, fiel die Menge der aufgenommenen Laute pro Säugling sehr unterschiedlich aus (nur spontane Lautproduktion).

Die aufgezeichneten akustischen Originalsequenzen enthalten neben den zu analysierenden Säuglingslauten auch unterschiedlich lange Pausen vor und während der Lautierungen, informative Ansagen sowie klinisch bedingte Nebengeräusche. Um die im Projekt zu analysierenden kindlichen Laute aus diesen Originalsequenzen zu separieren, sind Voranalysen nötig. Wie dazu vom Autor vorgegangen wurde, ist in Kapitel 3.6 beschrieben.

3.3 Einbettung der vorliegenden Arbeit in das Gesamtprojekt

Der Autor dieser Arbeit war im Rahmen der Kohortenstudie an der Aufarbeitung und Archivierung der Rohdaten beteiligt. Aufgrund des heterogenen Probandenkollektivs der Kohorte und aufgrund der Tatsache, dass es sich um die erste Kohortenstudie bzw. weltweit erste systematische Studie von Komfortlauten in diesem Alter handelt, erforderte die Auswertung neben einem hohen Grad an explorativen Ansätzen auch möglichst kontraststarke Gruppen, um geeignete Hypothesen für die nachfolgenden Untersuchungen der heterogenen Untergruppen sowie der Kohorte des Gesamtprojektes generieren zu können. Der Vergleich sollte in der vorliegenden Arbeit daher zwischen den Lauteigenschaften gesunder, normalhörender Probanden (NHS1-falsch-positive Gruppe; KG-Gruppe) und hochgradig hörauffälliger (HI-Gruppe) erfolgen. Als Kriterium für eine hochgradig sensorineurale Hörauffälligkeit wurde ein Hörverlust von $\geq 80\text{dB}$ verwendet.

In der Aufgabenstellung war zunächst nur die Einbeziehung von Säuglingen innerhalb der Kohortenpopulation vorgesehen. Aufgrund einer zu geringen Fallzahl zum Zeitpunkt der Erstellung vorliegender Arbeit aus der Kohorte wurden hier anonymisierte Daten weiterer geeigneter Fälle für die Kontrollgruppe aus vorhergehenden, methodisch vergleichbaren Studien des ZVES einbezogen. Auch die hörauffällige Gruppe wurde mit anonymisierten Daten von insgesamt vier Probanden vorhergehender methodisch vergleichbarer Längsschnitt-Studien am ZVES ergänzt. Dies verfolgte die Absicht in der vorliegenden Arbeit den Vergleich auf eine größere zugrundeliegende Datenbasis zu erweitern (vgl. 3.4).

Der Vorteil dieser Vorgehensweise bestand darin, dass die ergänzenden anonymisierten Aufnahmen aus dem Archiv im Rahmen einer Längsschnittstudie, also in zeitlich dichteren Abständen erhoben wurden, als die Kohortenstudien-Aufnahmen (Robb et al. 2020; Shehata-Dieler et al. 2013). Für die vorliegende Arbeit lieferten die Längsschnittdaten dieser Säuglinge auch eine Art Vergleichsmodell, da Referenzdaten in der Fachliteratur bisher komplett fehlen. Die spezifische Datenstruktur wurde durch Anwendung gemischter Modelle bei der statistischen Analyse berücksichtigt (vgl. Kap. 3.11). Aus dem ZVES-Archiv wurden vom Autor vorliegender Arbeit drei geeignete

Probanden und Probandinnen für die Kontrollgruppe sowie vier für die HI-Gruppe ausgewählt.

In Abbildung 1 ist der Rekrutierungsweg der Untersuchungsgruppen vereinfacht dargestellt. Die Untersuchungsgruppen vorliegender Arbeit aus der Kohorte sind grün markiert. Die ergänzenden Fälle aus dem Archiv sind blau/grau markiert.

Es wurden hier Komfortlauteigenschaften von 18 Säuglingen der KG-Gruppe mit denen von 8 Säuglingen der HI-Gruppe verglichen. Das Untersuchungskollektiv ist in Kapitel 3.4 charakterisiert.

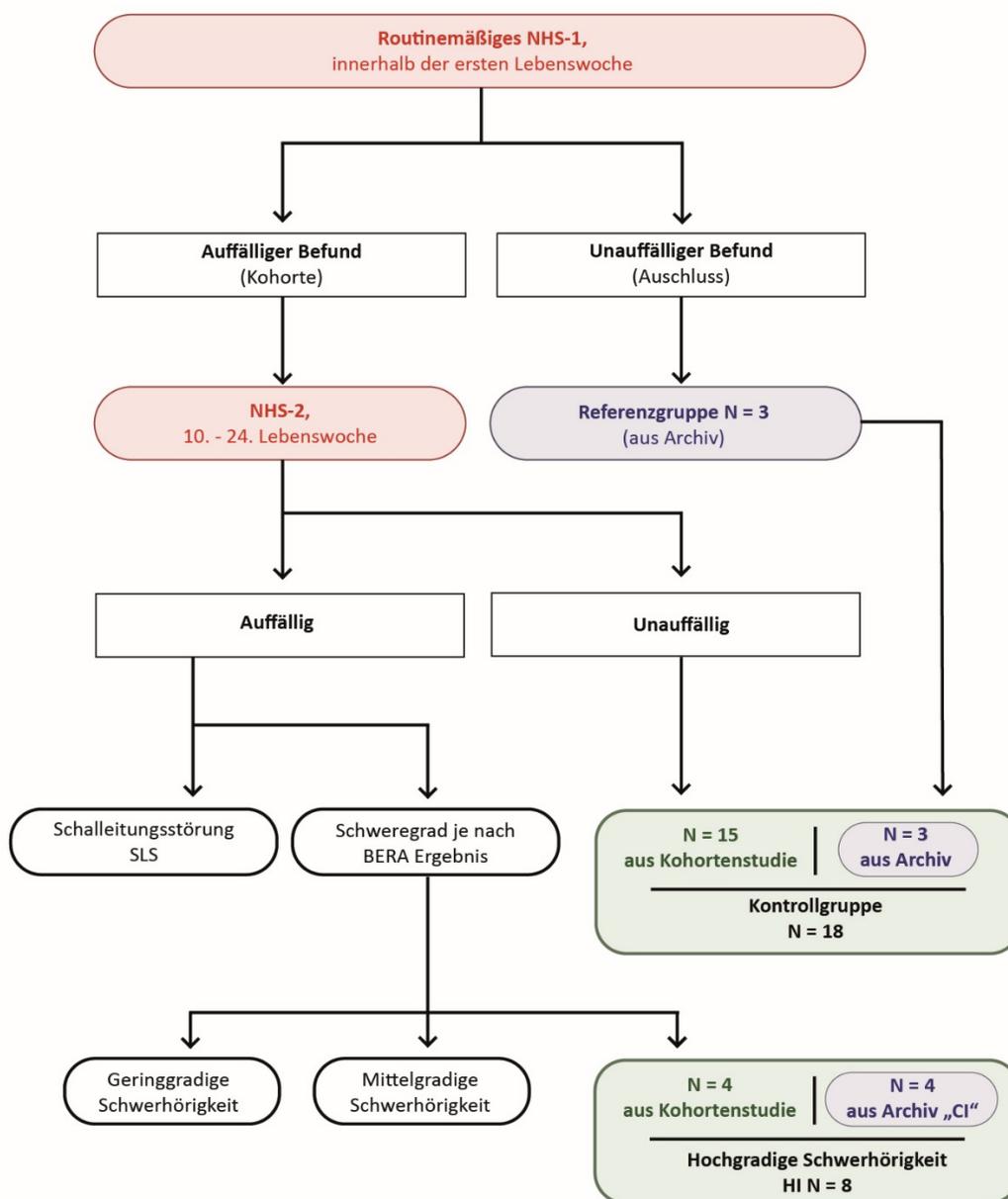


Abbildung 1: Rekrutierungswege der Untersuchungsgruppen (eigene Darstellung)

3.4 Untersuchungskollektiv

Nach vorgestelltem Rekrutierungsweg (vgl. 3.3) besteht das Untersuchungskollektiv aus drei Gruppen: Referenzgruppe, NHS-2 Kontrollgruppe und Gruppe „HI“ (aus aktueller Studie und Archiv). Die Kontroll- und Referenzgruppe, die hier gemeinsam die Gruppe der normalhörenden Probanden darstellen, bilden im Rahmen der Auswertungen gemeinsam die finale Kontrollgruppe.

3.4.1 Patientengruppe (HI)

Die Gruppe der Hörauffälligen (engl. hearing impaired / HI) umfasst Patienten und Patientinnen, die nach BERA-Screening beidseits Hördefizite von mindestens 80dB aufwiesen und nach WHO-Klassifikation als an Taubheit grenzend schwerhörig eingestuft werden.

Es wurden hier vier Säuglinge aus der Kohortenstudie (BA01-BA04) sowie zusätzlich vier aus dem Datenarchiv (BB01-BB04) eingeschlossen. Tabelle 1 enthält das Geschlecht sowie die somatischen Größen zur Geburt der Probanden und Probandinnen der HI-Gruppe.

Tabelle 1: HI-Gruppe (bilaterale hochgradige Hörstörung)

Säugling	Geschlecht	Perinataldaten		
		Gestationsalter in Wochen + Tagen	Gewicht [g]	Länge [cm]
BA01	♂	40	3.100	48
BA02	♀	41+3	3.960	51
BA03	♂	39+4	4.455	56
BA04	♂	41+2	3.810	50
BB01	♂	40	3420	53
BB02	♂	37+3	3.630	53
BB03	♀	34+5	2.315	46
BB04	♀	34+5	2.310	47

Während die Säuglinge aus der Kohorte zum Zeitpunkt der Aufnahmen alle noch unversorgt waren (BA01 - BA04), bekamen die Säuglinge der vorhergehenden Studien (Archiv; BB01 - BB04) zu verschiedenen Zeiten bereits Hörgeräteversorgungen im Untersuchungszeitraum der Arbeit (vgl. 3.5). Da es sich um hochgradige Hörminderung handelt ist kein Effekt durch die Hörgeräteversorgung auf die hier untersuchten Lauteigenschaften zu erwarten.

3.4.2 Kontrollgruppe

Für die Kontrollgruppe galten folgende Kriterien:

- Keine Schallempfindungsstörungen am Tag der Lautaufnahme (NHS-2) oder sonstige Höreinschränkungen im Verlauf des ersten Lebensjahres.
- Ausschluss mentaler, kardiovaskulärer, respiratorischer und syndromaler Grunderkrankungen.

Die Kontrollgruppe (KG) setzt sich aus insgesamt 15 Säuglingen der aktuellen Studie (AA01-AA15) sowie weiteren drei aus dem Archiv (AB01 – AB03) zusammen. Am Tag der Aufnahme (NHS-2) waren die Säuglinge der Kohorten-KG-Gruppe zwischen 63 und 149 Tagen alt. Von den drei Längsschnitt-Säuglingen aus dem Archiv wurden Längsschnittdaten im Alter von 65 bis 180 Tagen ausgewertet. Tabelle 2 enthält das Geschlecht sowie die somatischen Größen zur Geburt der Probanden und Probandinnen der Kontrollgruppe.

Tabelle 2: Kontrollgruppe

Säugling	Geschlecht	Perinataldaten		
		Gestationsalter in Wochen + Tagen	Gewicht [g]	Länge [cm]
AB01	♀	38+3	3.180	49
AB02	♀	39	3.075	49
AB03	♂	38	2.790	46
AA01	♀	41	3.120	51
AA02	♀	Keine Angabe	3.680	54
AA03	♀	39+2	3.470	51
AA04	♂	Keine Angabe		
AA05	♀	39+5	3.195	50
AA06	♂	38+5	4.180	53
AA07	♂	41+0	3.140	51
AA08	♂	38+6	3.910	51
AA09	♂	40+3	4.050	51
AA10	♀	38+3	3.420	51
AA11	♂	37+5	2.800	50
AA12	♂	40+0	4.610	54
AA13	♂	40+1	3.520	51
AA14	♀	37+1	2.850	50
AA15	♂	40+4	3.090	50

3.5 Übersicht der Aufnahmezeitpunkte und der Anzahl aufgezeichneter Laute

Die Lautaufnahmen in der Kohorte erfolgten im Rahmen des NHS-2 innerhalb der 10. bis 24. Lebenswoche (siehe Abbildung 1). In Tabelle 3 sind die Aufnahmezeitpunkte (Alter der Säuglinge in Tagen) an denen auswertbare Aufnahmen vorliegen, berichtet. Von den hier untersuchten Kohortenkindern lagen nur Aufnahmen bis zur 22. Lebenswoche vor. Von den Archivkindern beider Gruppen gab es Aufnahmen bis zur 26. Woche (Tabelle 3). Die in Tabelle 3 unterstrichenen Angaben markieren den Zeitpunkt der jeweiligen Hörgeräteversorgung bei den Säuglingen der HI-Gruppe aus dem Archiv.

Ab der 23. Lebenswoche gingen lediglich Aufnahmen der Archivkinder aus beiden Gruppen in die Auswertung ein.

Tabelle 3: Alter (Wochen) der Säuglinge bei Aufnahme, Anzahl aufgenommener Vokalisationen und Zeitpunkt der Hörgeräteversorgung (unterstrichene Tage) der Gruppe „HI“

Gruppe	Proband	Alter in Wochen																	
		≤10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
HI	BA01		72															128	
	BA02				88														
	BA03								111										
	BA04			79															
	BB01							<u>100</u> 103		117					156 160	163	169		
	BB02				85	91		98											
					87	93		100	109	112	120		139			156		171	181
					88	94		100	109	112	120		139			156		171	181
				<u>89</u> 90	95		102												
	BB03			79				99				127	141		<u>156</u>		171		
BB04			79				99				127	141		<u>156</u>		171			
KG	AA01							109											
	AA02		63																
	AA03									113									
	AA04								107										
	AA05									113									
	AA06										125								
	AA07								106										
	AA08		60																
	AA09											131							
	AA10			82															
	AA11													149					
	AA12		65																
	AA13										124								
	AA14											134							
	AA15							99											
AB01		65		87		104				120		134 139	141 142 146	148	154 155	162 163	171	180	
AB02			80		94		105		121				142 143	149 150	154 156 157 159	161 164 167	173		
AB03			82		96		105				126		145	148	154 159	167		179	

3.6 Aufbereitung und Analyse der Komfortvokalisationen

Abbildung 2 zeigt den schematischen Ablauf der umfassenden Datenaufbereitung und -auswertung bis hin zur statistischen Auswertung. Die Darstellung dient der Übersicht und chronologischen Einordnung der im Folgenden detailliert beschriebenen Methoden in den Kapiteln 3.6.1 bis 3.11.

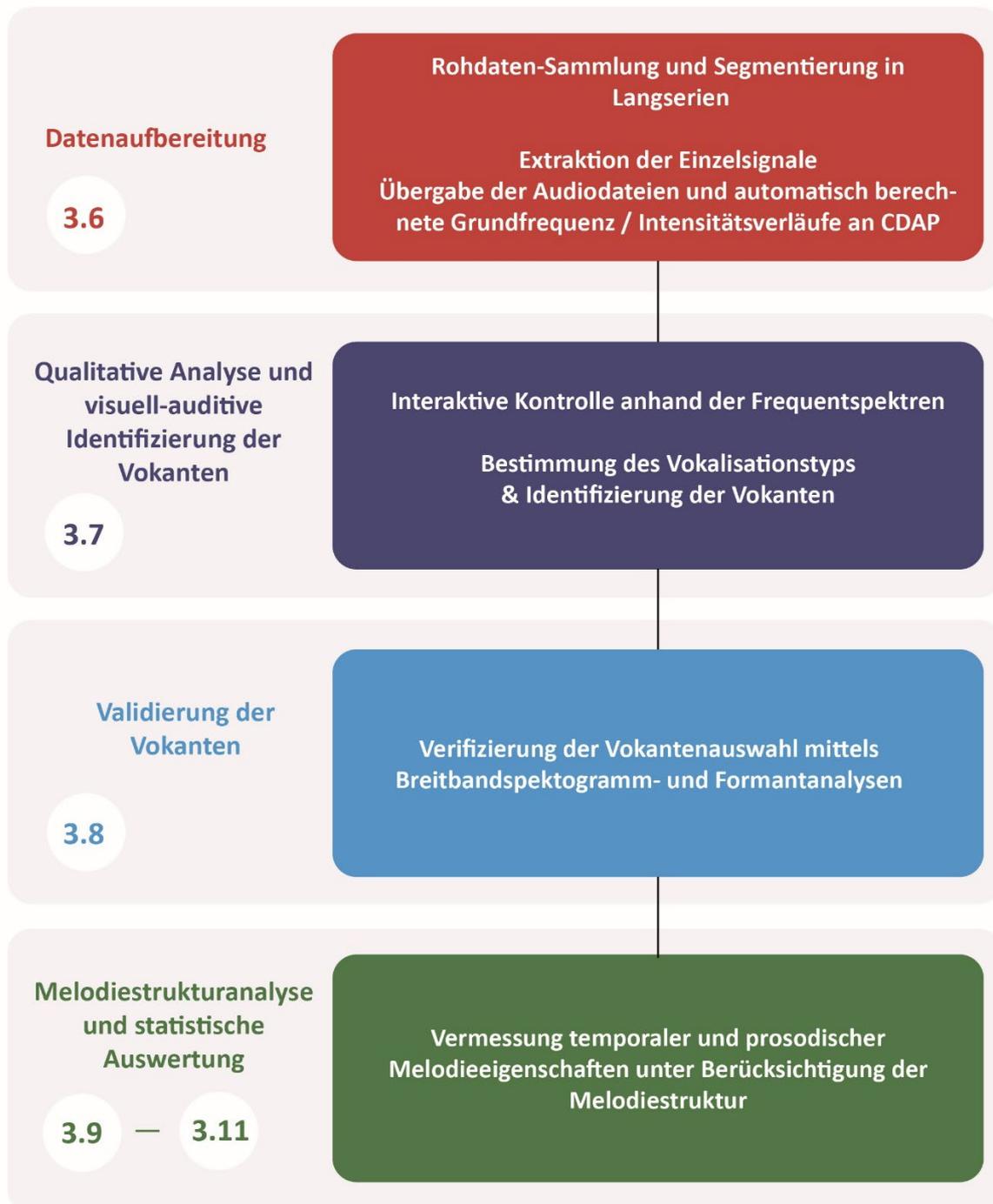


Abbildung 2: Prozess der Datenverarbeitung [Erklärungen in den nachfolgenden Kapiteln]

3.6.1 Von der Audiodatei zur Langserie (Datenaufbereitung)

Die am CHC digital aufgenommenen Rohdaten (Sequenzen spontaner Lautproduktionen) (Kap. 3.2) wurden pseudonymisiert an das ZVES zur Aufbereitung und spezifischen Auswertung weitergegeben. Zunächst wurden die digital erfassten Original-Audiodateien (WAV-Files) zur besseren Handhabung bei der weiteren Analyse in kürzere Abschnitte, in sogenannte Langserien, mithilfe der open-source Software PRAAT (Paul Boersma & David Weenink 2018) geschnitten und archiviert. Dabei wurde bereits eine erste Selektion nicht brauchbarer Zeitabschnitte vorgenommen, in denen keine Komfortlaute des Säuglings vorkamen oder diese durch Hintergrundgeräusche oder andere Störfaktoren überlagert waren. Dieses Vorgehen vereinfachte nicht nur die Bearbeitung der folgenden Schritte, sondern lieferte auch einen ersten Überblick des gesammelten Datenmaterials. So konnten Aufnahmeeinheiten ggf. bereits hier, aufgrund mangelnder verwendbarer Sequenzeinheiten aussortiert werden.

3.6.2 Von der Langserie zum Einzelsignal: Schnitt der Langserien und Segmentierung der Einzelereignisse

Anschließend erfolgte die Annotation der Langserien mittels PRAAT (s. Kap. 3.6.3.) und damit die Klassifizierung der Einzelereignisse. Einzelereignisse sind Vokalisationen des Säuglings, Atemgeräusche und Folgen von Kurzlauten sowie weitere in Tabelle 4 aufgeführte Ereignisse.

Tabelle 4: Annotationsvariablen für Einzelereignisse

e	Phonation während einer Exspirationsphase
i	Inspirationsgeräusch
a	schnelle akustisch wahrnehmbare Atemsequenzen ohne Lautierung
f	„fuzzing“; sehr kurze Weinlaute (< 100ms), oft in Folge hintereinander
in	Interaktionsserie; vokale Stimulation durch die Kontaktperson mit nachfolgend erkennbarer und direkter stimmlicher Reaktion des Säuglings
kp	Äußerungen der Kontaktperson separat mit zeitlich versetzter Reaktion des Säuglings
nv	„nicht verwendbar“ - Ansagen, rein vegetative Begleitgeräusche, Weinen Störgeräusche etc.

In dieser Arbeit wurden lediglich die Zeitintervalle mit der Annotationsvariablen „e“ und damit die Phonationen von Komfortvokalisationen während der Expirationsphase untersucht. Die vollständige Annotation der Langserien diente jedoch dem übergeordneten Projekt und stellte eine inhaltliche Einordnung des analysierten Datenmaterials für Nachfolgestudien dar.

Ein im ZVES erstelltes PRAAT-Skript, geschrieben und zur Verfügung gestellt von Michael Eisenmann (Eisenmann), half, durch automatisierte Cursor-Setzung nach Abtasten der Intensitätseigenschaften, eine erste Differenzierung zwischen Expirations- und Inspirationsphasen innerhalb der Langserien zu ermöglichen.

Vor der automatischen Extraktion der hier analysierten egressiven Vokalisationen (e) wurde jeder automatisch annotierte Abschnitt nochmals vom Autor kontrolliert, durch weitere Variablen (Tabelle 4) ergänzt und ggf. Cursorpositionen für den Beginn und das Ende des Lautes manuell nachkorrigiert, entfernt oder ergänzt.

3.6.3 Extraktion der egressiven Komfortvokalisationen („e“) und Übergabe der Audiodateien und f_0 -Werte an CDAP

Im letzten Schritt konnte der eigentliche Extraktionsvorgang der Audiodateien der final gewählten egressiven Phonationen mithilfe eines weiteren Skripts, dem sogenannten WAV-Extraktor (Eisenmann), vorgenommen werden. Hierdurch konnten alle Audiodateien der mit „e“ annotierten Abschnitte (e; Tabelle 4) innerhalb der Langserie automatisch ausgelesen und in ein vom Autor gewähltes Verzeichnis am Computer exportiert werden. Bei den mit „e“ annotierten Komfortvokalisationen handelte es sich jedoch nicht ausschließlich um Vokanten. Vielmehr wurde ein übergreifendes Komfortlautrepertoire zusammengestellt, das den Untersuchungen im Rahmen der gesamten Kohortenstudie dienlich sein würde. Differenziert und klassifiziert wurden hierbei vom Autor also auch komplexere Lautbildungen, wie erste supralaryngeale Konsonantbildungen und andere artikulatorische Elemente (vgl. Tabelle 5).

Für jeden egressiven Laut (e) wird zusätzlich eine automatische Grundfrequenz-/ Intensitätsanalyse durchgeführt. Dazu wurde ein weiteres Praat-Skript verwendet, das

auf einer Autokorrelationsmethode basiert (Eisenmann). Alle Audio- und Frequenz-/Intensitätsdaten wurden dann automatisch an das laboreigene CDAP-System (Programmator: P. Wermke – „PW-project“) übergeben.

In Abbildung 3 ist ein PRAAT-Ausgabe-Fenster beispielhaft abgebildet.

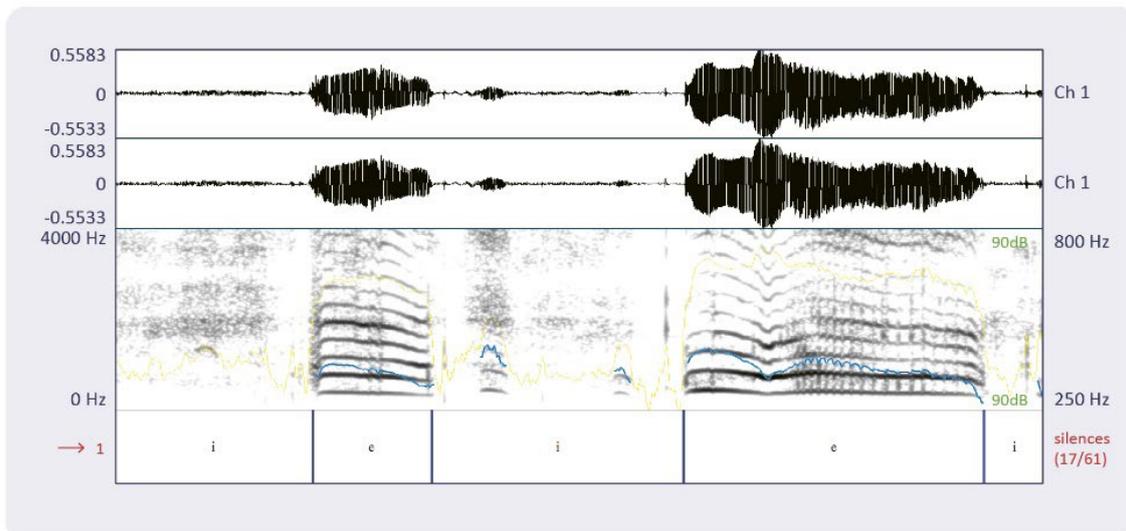


Abbildung 3: Beispiel einer annotierten Aufnahmesequenz im PRAAT-Ausgabefenster

Im oberen Abschnitt des Ausgabefensters ist die Regleramplitude der Sequenz dargestellt. Im mittleren Abschnitt das Frequenz-Spektrogramm (Frequenzbereich hier: linear 0-4000Hz) und der Grundfrequenzverlauf (blau) ist dem Spektrum überlagert abgebildet. Die Grundfrequenz ist hier im Bereich zwischen 250 und 800Hz gezeichnet (Skala rechts). Der untere Bereich des Ausgabefensters zeigt die Annotationsvariablen der Einzelereignisse.

Hier besteht die Sequenz aus einem Inspirationsgeräusch (i), gefolgt von einem Vokanten (e) mit einbögiger Melodie, dem wieder eine Inspiration folgt. Der zweite egressive Laut (e) zeigt eine zweibögige Melodie und ist kein Vokant, da er eine velare Konsonantenbildung (Artikulation) enthält. Am Ende gibt es ein weiteres Inspirationssignal (i).

3.7 Qualitative Analyse und audiovisuelle Identifizierung der Vokanten

Die extrahierten Einzelereignisse wurden als wav-files (Audio-Dateien) an das im ZVES verwendete CDAP (Cry-Data-Analysis-Program; Programm-Autor P. Wermke) übergeben und für die weitere Bearbeitung bereitgestellt. Auch die in PRAAT ermittelten Grundfrequenzwerte (f_0) bzw. die Werte ihrer Zeitfunktion (Melodie) und die Frequenzspektren wurden an CDAP übergeben.

3.7.1 Interaktive f_0 -Kontrolle

Zur Kontrolle aller mittels PRAAT automatisch berechneten f_0 -/Intensitätswerte wurden diese im CDAP-Modul „Melodie-/Intensitätsdiagramm“ (EFA Modus) geeignet grafisch dargestellt. Zusätzlich wurde das zugehörige Frequenzspektrum der jeweiligen egressiven Vokalisation (e) mit abgebildet. Die Frequenzspektren der einzelnen Audiodateien wurden dafür routinemäßig mithilfe des Kay CSL 4300; Computerized Speech Lab (Bahgat et al. 2012) am ZVES berechnet und in das CDAP-System als Bildobjekte eingefügt (EFA-Modus und Albenmodus). Über diese können Melodie- und Intensitätsdiagramme der zugehörigen Frequenzdaten auf diese Bildebene projiziert und anschließend manuell vom Bearbeiter auf Richtigkeit überprüft werden.

Es wird dabei untersucht, ob der mit PRAAT automatisch berechnete Melodieverlauf (Zeitfunktion von f_0) korrekt ist bzw. sich, im Fall von Subharmonischen oder Oktavsprüngen im richtigen Frequenzbereich befindet (Ausschluss von Berechnungsartefakten). Einen Anhalt dafür liefert das Schmalbandspektrogramm, welches die Frequenzbänder bis 4kHz im 500Hz Grid angibt und somit eine grobe Abschätzung der Grundfrequenz und ihres Verlaufs anhand der höheren Harmonischen zulässt. Die Grundfrequenz entspricht dabei der Schwingung der Stimmlippen, während es sich bei den darüberliegenden Frequenzbändern im Spektrum um Obertöne handelt, die ein ganzzahliges Vielfaches von f_0 darstellen und durch Resonanzen im Vokaltrakt beeinflusst werden.

Bei offensichtlichen Fehlern der automatisch ermittelten f_0 -Kontur im Vergleich mit dem Spektrum erfolgte eine manuelle Neuberechnung; die entsprechenden Audiofiles werden durch einen Befehl in CDAP der PRAAT-Kontrolle zugeteilt. Nachfolgend wird

in PRAAT durch Wahl neuer, manuell gewählter Einstellungen der korrekte f_0 -Verlauf ermittelt und final an CDAP zurückgegeben. Diese Kontrolle ist sehr zeitaufwendig, für eine zuverlässige weitere Analyse der Lautmelodien aber unerlässlich.

In einem weiteren Analyseschritt werden alle f_0 -Verläufe (Melodie) in CDAP gefiltert (40Hz-Lowpassfilter) und Extremwerte gelöscht, um sie für eine finale Vermessung der Melodieeigenschaften zu speichern (MMD-Modus).

3.7.2 Bestimmung des Vokalisationstyps

Im Altersbereich der Untersuchung zeigen Komfortlaute (Laute in positiver Grundstimmung) ein Repertoire verschiedenster Lautarten (Vokalisationstypen). Dieses Repertoire soll im übergeordneten Gesamtprojekt (Kohortenstudie) möglichst detailliert beschrieben und zwischen den Säuglingen in Abhängigkeit vom Hörstatus verglichen werden.

In der vorliegenden Arbeit wurde nur der „einfachste“ Vokalisationstyp des Komfortlautrepertoires, die sogenannten Vokanten untersucht. Der Begriff Vokant (engl. vocant) stammt von Martin (Martin 1981). Er bezeichnet sie als vokalähnliche Phonationen ohne intentionale artikulatorische Aktivität, also mit neutraler Position des Vokaltrakts.

Da Vokanten im Gegensatz zu anderen Vokalisationstypen (Tabelle 5) vergleichsweise eindeutig via audiovisueller Analyse identifiziert werden können (Buder et al. 2013) und bei allen Säuglingen neben dem Weinen als früher Vokalisationstyp auftreten, eignen sie sich sehr gut für einen Vergleich zwischen der Kontroll- und Hörgestörten-Gruppe.

Wie in der Einleitung erläutert, erfolgt die Klangerzeugung bei den Vokanten laryngeal (Phonation) bzw. teilweise zusätzlich durch laryngeale Konstriktionen (prä-artikulatorische Aktivität), während der Vokaltrakt unverändert bleibt („neutrale“ Position), sodass die vokaltypischen Resonanzen (Formanten) nicht auftreten. Vokanten sind damit nur Vorstufen der späteren Vokale. Buder et al. nennen sie „Quasivokale“ („Quasivowels“) im Gegensatz zu vollen Vokalen („full vowels“). Oller hat den Begriff der „Quasivowels“ eingeführt und zählt sie gemeinsam mit anderen Lauten zu den von

ihm definierten „Protophonen“ (Oller und Kent 2001). Vokanten mit Quasivokalen bilden die einfachste und früheste Form der Nichtschrei-Vokalisationen oder Komfortlaute.

Die Identifizierung der hier untersuchten Vokanten erfolgte neben audiovisuellen Analysen (Tabelle 5) unter Verwendung von Schmalbandspektogrammen und Breitbandspektogrammen sowie Formantanalysen mittels LPC (Linear Predictive Coding).

Vokanten zeigen sehr häufig am Anfang oder am Ende laryngeale Konstriktionsphänomene (Buder et al. 2013), die nach dem „Laryngeal Articulator Model“ (LAM-Modell) von Esling (Esling 2005) vor allem durch Verengungen der aryepiglottischen Falten zustande kommen (vgl. Kap. 1). Ein Resultat solcher laryngealen Aktivitäten sind kurzzeitige Konstriktionen am Vokantenanfang und/oder -ende oder vollständige Glottisverschlüsse, sogenannte Segmentierungspausen (SEG) (Yavarzadeh 2020). Diese Phänomene stehen am Beginn der artikulatorischen Entwicklung noch bevor supralaryngeale artikulatorische Aktivitäten auftreten (Esling 2005; Gregory et al. 2018). Vokanten können auch in Kombinationen mit Vokalen (full vowels) und anderen artikulatorischen Elementen (Konsonanten) vorkommen. Aus Gründen der Homogenität wurden hier solche Lautäußerungen ausgeschlossen (Blum; Buder et al. 2013). In der vorliegenden Arbeit wurden nur einfache, „isoliert“ auftretende Vokanten analysiert.

Tabelle 5: Vokanten im Vergleich zu anderen beobachteten Vokalisationstypen im analysierten Datenmaterial

Ursprung der Klangbildung	Vokalisationstyp	Beschreibung
Larynx (Phonation) und laryngeale Konstriktion	Vokanten (vocants)	vokalähnliche egressive Laute ohne supralaryngeal erzeugte artikulatorische Eigenschaften (Formanten) Oszillationspausen der Stimmlippen und laryngeale Konstriktionen können vorkommen
	sehr hochfrequente Laute (squeals)	Laute in denen die Grundfrequenzwerte über der normalen Tonlage der Säuglinge liegen und als hochfrequentes Quieten wahrgenommen werden. Frequenzen über 700Hz werden über weite Bereiche oder im Mittel über den gesamten Laut erreicht („High Pitch“). Laryngeale Konstriktionen sind hierbei typisch.
Vokaltrakt (supra-laryngeale Artikulation)	„Raspberries“ (Vibrationslaute)	sogenannte Raspberries treten bei erhöhtem intraoralem Druck und gleichzeitig entspannter Lippenhaltung auf. Je nach Stellung der Zunge, ob anterior zwischen den Lippen oder im entspannten Zustand auf dem Mundboden liegend, entstehen dabei unterschiedliche Klangeindrücke (Vibrationen). Neben der akustisch mühelosen Identifizierung zeichnen sich Raspberries durch ein typisches Muster im Spektrum aus. Über das gesamte Frequenzspektrum zeigen sich vertikale „Streifen“, die die Melodie überlagern.
	Closants & Konsonanten	konsonantähnliche Elemente und Konsonanten, die in Kombination mit Vokanten und Vokalen auftreten (cv)
	Volle Vokale	intentionale Änderungen der Resonanzeigenschaften des Vokaltrakts, die durch Formantfrequenzbewegung identifiziert wurden

3.8 Validierung der Vokanten

3.8.1 Breitbandspektogramm-Analyse

Die Identifikation des hier analysierten Vokantentyps erfolgte zunächst audiovisuell unter Verwendung der Frequenz-Spektrogramme und Audiofiles. Die Vokanten, die hier untersucht wurden, kann man sehr gut anhand ihrer akustischen Eigenschaften per Gehör identifizieren (Buder et al. 2013). Sie werden mit einer „neutralen“ Vokaltrakt-Stellung erzeugt und klingen manchmal wie der Schwa-Laut.¹ Um zusätzlich eine objektive Methode anzuwenden, die die audiovisuelle Identifikation ergänzt, wurden für alle audiovisuell identifizierten Vokanten neben Breitbandspektogrammen (Abbildung 4) auch Formanten berechnet.

Mithilfe der Breitbandspektogramme können Resonanzeigenschaften des Vokaltraktes abgebildet werden, durch Formantanalysen werden diese objektiviert.

Von einer Gesamtzahl von N=1369 anhand der Audiofiles und Schmalbandspektogramm vorausgewählten Vokanten, konnten final N=1236 mithilfe der Breitbandspektogramme als einfache Vokanten identifiziert werden (Tabelle 6). Diese Zahl setzt sich aus 588 Lauten bei der hörgestörten Gruppe und 648 Lauten bei der Kontrollgruppe zusammen.

Die Rohdatenbasis der Auswertung bildeten alle egressiven Vokalisationen der Aufnahmen im Untersuchungszeitraum pro Säugling, die innerhalb der Kohortenstudie aufgezeichnet wurden bzw. bereits im Archiv vorlagen (Kap. 3.4). Da die Säuglinge Komfortlaute spontan erzeugen, je nach affektivem Zustand und „Lautierlaune“, unterscheidet sich die Anzahl der aufgezeichneten Vokalisationen zwischen den Säuglingen.

¹ Das e-Schwa (<e>) steht für den mittleren Zentralvokal. Im Deutschen erscheint er nur in unbetonten Silben (Staffeldt (2010))

In Tabelle 6 sind die Aufnahmezeitpunkte und die an diesen aufgezeichneten Vokalisationen berichtet, sowie die identifizierte Anzahl der Vokanten (vgl. 3.7.2) für alle individuellen Säuglinge beider Gruppen zusammengefasst.

Tabelle 6: Egressive Vokalisationen gesamt und Anteil Vokanten der Kontroll- (KG) und hörgestörten Gruppe (HI)

Gruppe	Säugling	Aufnahme- termine	Alle Laute	Vokanten absolut	Vokanten relativ	
HI	BA01	2	125	20	16%	
	BA02	1	46	32	70%	
	BA03	1	157	37	24%	
	BA04	1	169	113	67%	
	BB01	6	345	39	12%	
	BB02	21	652	110	17%	
	BB03	6	773	179	23%	
	BB04	6	532	58	11%	
		ΣHI	45	2847	588	Ø= 20,8%
KG	AA01	1	52	19	37%	
	AA02	1	20	3	15%	
	AA03	1	12	3	25%	
	AA04	1	21	2	10%	
	AA05	1	66	11	17%	
	AA06	1	92	43	47%	
	AA07	1	114	11	10%	
	AA08	1	14	1	7%	
	AA09	1	57	7	12%	
	AA10	1	29	5	17%	
	AA11	1	62	10	16%	
	AA12	1	68	14	21%	
	AA13	1	46	8	17%	
	AA14	1	73	18	25%	
	AA15	1	21	7	33%	
	AB01	16	1100	193	18%	
	AB02	16	1163	160	14%	
	AB03	10	1122	133	12%	
		ΣKG	57	4151	648	Ø= 15,7%

Abbildung 4 zeigt Beispiele von Breitbandspektrogrammen zweier typischer Vokanten. Zu sehen sind farblich nach Energieintensität codierte starre Formanten. Frequenzbereiche mit niedriger Intensität sind braun, hohe Energiedichten dagegen weiß bzw. lila codiert. Entlang der Abszisse handelt es sich um eine lineare Skalierung der Frequenzen von 0 bis zu 4 kHz über der normierten Zeitachse (Ordinate). Die Darstellung entspricht einer Bandbreite von 323Hz (Hanning Window).

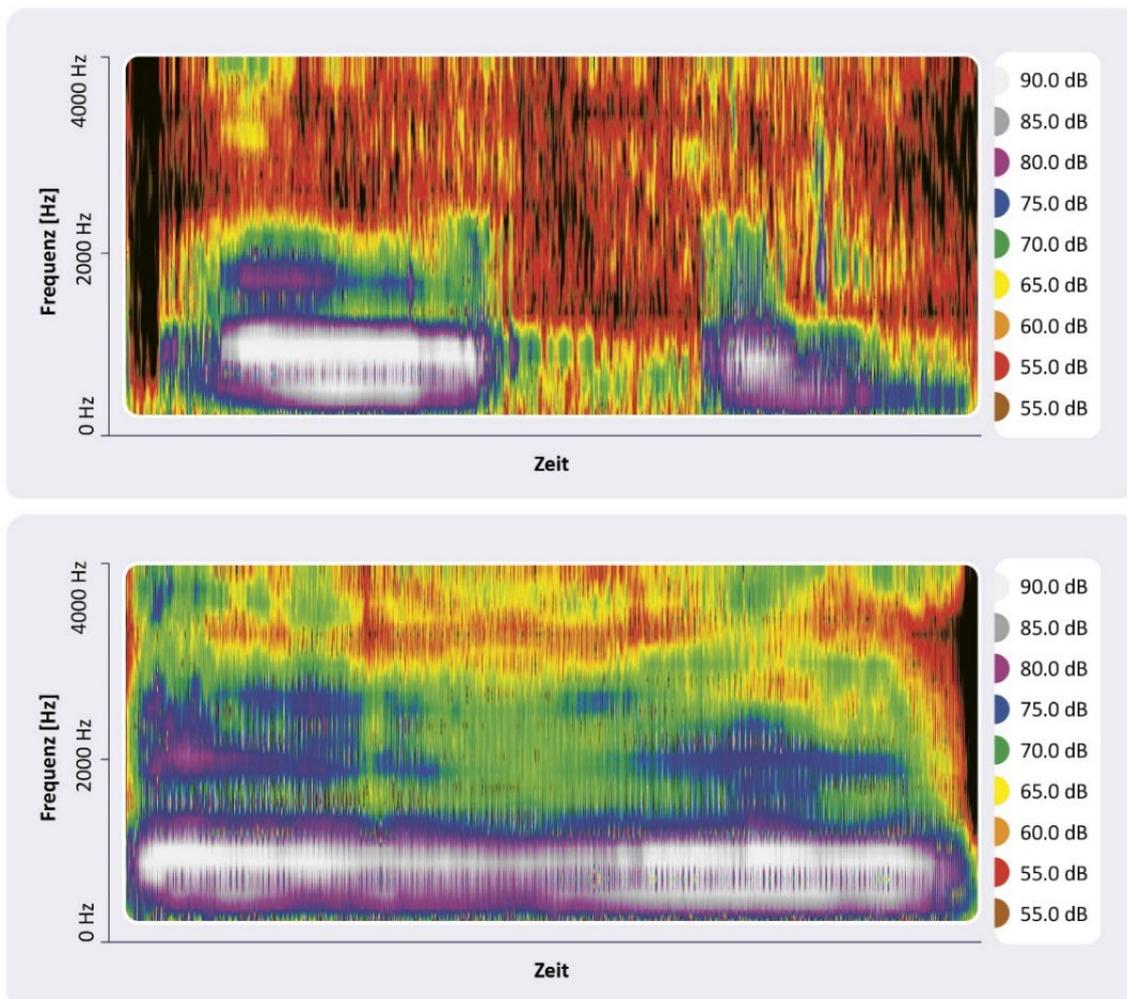


Abbildung 4: Breitbandspektrogramme zweier Vokanten

Die tiefsten Resonanzbänder (Formant F1) lassen sich gut entlang der weißen horizontalen Basisstränge identifizieren und weisen die höchsten Energiedichten bei ungefähr 730Hz bzw. 750Hz auf. Die zweite Resonanzfrequenz (Formant 2) ist aufgrund niedrigerer Energieniveaus visuell im Spektrogramm oftmals schwerer identifizierbar. Die Frequenzwerte für F2 liegen in diesen Beispielen bei ungefähr 1570Hz bzw. 1800Hz. Beide Abbildungen entsprechen dem typischen Formantbild eines Vokanten. Die artikulatorische Inaktivität wird durch die parallel verlaufenden, keine Bewegung zeigenden, Resonanzbanden entlang eines sehr schmalen Frequenzbereiches deutlich und mittels farblicher Spektrographie hier veranschaulicht.

3.8.2 Formantanalyse

Bei der laryngealen Lautproduktion (Phonation) wird die akustische Energie, die durch die Stimmlippen-Schwingung erzeugt wird, durch die geometrischen und intentionalen Vokaltrakt-Formveränderungen in Resonanz gebracht. Die Vokaltrakt-Resonanzen konzentrieren die akustische Energie in diskreten Frequenzbändern, die man in der Linguistik Formanten nennt. Je kürzer der Vokaltrakt, desto höher die Formantfrequenzen. Für die Wahrnehmung der Vokalqualität in der Sprache sind vor allem der erste (F1) und der zweite (F2) Formant wichtig. Vokanten sind nur vokalähnlich und zeigen damit noch nicht die intentional erzeugten F1 und F2 Werte der Vokale in der Sprache. In den Breitbandspektogrammen kann man die Formantfrequenzen grob abschätzen. Wichtig für die Identifizierung von Vokanten war es, dass sich die Formantfrequenzen über den Laut nicht verändern, also keine Artikulation erfolgt.

Um die korrekte Auswahl der Vokanten im Lautrepertoire vor ihrer weiteren Analyse zusätzlich zu den auditiven Analysen und Analysen mittels Breitbandspektogramm zu objektivieren, wurden die ersten beiden Formantwerte (F1 und F2) bestimmt. Die Bereiche für die Formantfrequenzen, die bei Vokanten die Resonanzeigenschaft des Vokaltraktes in neutraler Position reflektieren, wurden anhand der Variationsbereiche für Formanten im Alter von 4-6 Monaten aus den Analysen von Masapollo (2016) definiert. Hierbei galten für die automatische Berechnung der Formantdaten folgende Vorgaben:

- (1) Frequenzbereich Formant 1: 500-1700Hz
- (2) Frequenzbereich Formant 2: 1100Hz-3500Hz

Zur automatischen Formantanalyse wurde hier ebenfalls das CSL-4400 von Kay Pentax verwendet. Sie erfolgte mittels „Linear Predictive Coding (LPC)-Algorithmus“ (Adaptive Autokorrelationsmethode; 30 Filterungen; Hanning window; Preemphasize: 0,9; Fensterlänge: 10ms; Schrittweite: 5ms). LPC ist ein Standardverfahren zur Formantanalyse in der Linguistik. Formantanalysen in Säuglingslauten sind schwierig,

da die Grundfrequenzen höher sind als in der späteren Sprache (Kempf 2010; Robb et al. 1994; Robb und Cacace 1995). Hohe Grundfrequenzen (f_0) können zusätzlich die Berechnung erschweren. Geraten sie in den Bereich des ersten Formanten (F1), erzeugen diese Fehler bei der automatischen Formantanalyse (Bauer und Kent 1987; Robb und Cacace 1995; Wermke 2002). Daher wurden in der vorliegenden Arbeit die Formantfrequenzen nur für Vokanten mit einer Grundfrequenz von maximal 400Hz bestimmt und ausgewertet, um die Unsicherheiten des LPC-Algorithmus bei höheren Grundfrequenzwerten zu vermeiden (Kempf 2010; Wermke et al. 2002). In Einzelfällen wurden interaktiv Spektralscheiben mit unterschiedlicher Fensterlänge erstellt und dadurch eine zusätzliche Kontrolle ermöglicht.

Fielen über 50% der ermittelten einzelnen F1 und F2 Frequenzdaten pro Laut außerhalb des in (1) und (2) definierten Frequenzbereichs, wurde dieser bei der Darstellung im Vokaldreieck nicht berücksichtigt. Das wird dadurch begründet, dass man in diesen Fällen nicht von stabilen Resonanzen, also artikulatorischer Inaktivität, ausgehen konnte. Bei der hörgestörten Gruppe konnten bei 331 Vokanten (56%), bei der Kontrollgruppe bei 352 (54%) der Vokanten F1 und F2-Werte ermittelt werden.

Es zeigten sich bis 10kHz drei Hauptresonanzbänder F1-F3. In der vorliegenden Arbeit wurden nur die ersten zwei Formanten im Bereich bis zu 3,5kHz berechnet, da diese die typischen Vokalformanten in der Sprache sind, die hier zur Validierung der selektierten Vokanten verwendet wurden. Um zu zeigen, dass die hier getroffene Auswahl bezüglich der Formantfrequenzen im Analysealter auch repräsentativ ist, wurden im Ergebnisteil (vgl. Kap. 4.1) die Formantfrequenzen in einem sogenannten Vokaldreieck dargestellt. Traditionell werden F1 und F2 in einem Diagramm so dargestellt, dass sich die F1-Werte auf der Y-Achse und die F2-Werte auf der X-Achse befinden. Für jeden Vokanten gibt es einen Wert in diesem Diagramm². Alle Punkte zusammen reflektieren

² Bei Untersuchung von zwei a-Vokalen spricht man auch vom Vokaltrapez.

den akustischen Vokalraum. Das Vokaldreieck wird durch die Formant-Werte der Vokale /i/, /a/ und /u/ aufgespannt und beinhaltet alle theoretisch möglichen Vokale für eine gegebene Vokaltrakt-Länge. Damit unterscheiden sich die Vokaldreiecke von Säuglingen, Kindern und Erwachsenen. In der Literatur gibt es relativ wenige Berichte über Formanteigenschaften im Analysealter der Arbeit, was mit den vorab genannten Schwierigkeiten in der Formantbestimmung bei hohen Grundfrequenzen zusammenhängt. Die f_0 von Frauen- und Kindstimmen liegt zwischen 196-262Hz bei Frauen bzw. 220-660Hz bei 6-jährigen Kindern, die von Säuglingen ist noch höher (Biesalski 1994, S. 277–281). Hier wurde das Vokaldreieck von Masapollo als Referenzgröße für die eigenen Werte herangezogen (Masapollo et al. 2016).

Bei den Vokaldreiecken werden nicht die Hertz-Angaben, sondern eine andere Messgröße, nämlich Bark dargestellt. Die Frequenzachsen sind in Bark-Einheiten skaliert, die die Formant-Frequenzen in Einheiten gleicher psychoakustischer Distanzen transformieren. Die BARK-Skala basiert auf der Einteilung des hörbaren Frequenzbereichs in 24 Frequenzgruppen (Delb und Praetorius 2004).

Es handelt sich dabei, um eine psychoakustische Größe, die ausgehend von den physikalisch gemessenen Frequenzwerten nach Umrechnung den physiologisch-auditorisch wahrgenommenen Klangeindruck widerspiegelt. Die in der Arbeit ermittelten Formant-Frequenzen wurden daher mittels Formel (1) in Bark-Werte vom Autor umgerechnet und im Ergebnisteil verwendet.

$$z = 13 \arctan(0,00076 * f) + 3,5 \arctan\left(\frac{f}{7500}\right)^2 \quad (1)$$

f = Grundfrequenz [Hz]
 z = Barkwert

Formel entnommen aus (Fruhmann); [02.05.2021.; 16.26]

3.9 Melodiestrukturanalyse einfacher Vokanten

Die Arbeit von Blum hat gezeigt, dass Vokanten unterschiedliche Melodiemuster enthalten können. Die meisten Vokanten haben eine einfache steigend-fallende Melodie (einböigige Muster). Es gibt aber auch Vokanten mit komplexen Melodiemustern, die aus mehreren Bögen bestehen (Blum).

Die Strukturanalysen der Melodie wurden mithilfe des ZVES internen Softwaresystems CDAP (Programm Autor P. Wermke) durchgeführt. In diesem Programm werden die mittels Praat berechneten Grundfrequenzverläufe (Kap. 3.6) logarithmisch gemeinsam mit der Intensität in Melodiediagrammen dargestellt.

Es lassen sich Melodiestrukturkategorien basierend auf der Bestimmung der Melodiebogenanzahl und Pausen objektiv bilden (Wermke 2002). Es wurde hier eine Kurzform der Melodiestrukturanalyse, abweichend von der am ZVES routinemäßig angewandten, verwendet.

Melodiebögen beschreiben einen in der Regel, mit Ausnahme von innermelodischen Pausen, unterbrechungsfreien f_0 -Konturverlauf (Ausnahme kurze Elemente mit phonatorischem Rauschen). Ob es sich um eine einfache oder mehrböigige Melodie handelt, wurde für jeden Vokanten bestimmt. Im Rahmen der Klassifizierung wurde ein Melodiebogen durch einen Frequenzhub zwischen f_0 -Minima und f_0 -Maxima von zwei Halbtonschritten auf der Seite der längeren Flanke definiert.

Folgende Melodiestrukturkategorien wurden hier für die Vokanten unterschieden:

3.9.1 1B und Kurzlaute (einfache Struktur)

Vokanten, deren Melodie aus einem einzelnen Melodiebogen (1B) bestand, wurden von Melodien mit einer Dauer von unter 300ms (Kurzlaute) unterschieden. Das Kriterium von 300ms findet sich allgemein in der Fachliteratur als Cut-Off-Kriterium (Lind 2007). Kurzlaute wurden hier von rein vegetativen Begleitgeräuschen durch auditive Analyse sowie ein weiteres Zeitkriterium separiert. Sie mussten eine Mindestlänge von 150ms haben. Das hat das Ziel, dass man vegetative Begleitgeräusche physiologischer Ereignisse (z.B. Schlucken, Husten) von intendierten kurzen Vokanten trennen möchte.

Hier wurden nur Vokanten-Kurzlaute einbezogen, bei denen durch auditive Kontrolle und Informationen aus den Begleitprotokollen zu den Aufnahmen sicher war, dass es sich nicht um rein vegetative Begleitgeräusche gehandelt hat.

3.9.2 Doppelbögen (2B) und segmentierte Strukturen (SEG) (Komplexe Strukturen)

Bei den Doppelbögen handelt es sich um Melodien, die aus zwei aufeinanderfolgenden auf- und absteigenden Grundfrequenzkonturen (2 Bögen - 2B) bestehen. Da auch noch komplexere Melodien mit mehr Bögen in Vokanten beobachtet wurden, hat der Autor eine weitere komplexe Gruppe gebildet, die aus komplexen Melodien mit mehr als zwei Bögen besteht (>2B). Aufeinanderfolgende Bögen können aufgrund laryngealer Konstriktionen Pausen, also sogenannte rhythmische Segmentierungen (SEG) aufweisen. Segmentierungspausen wurden hier als Unterbrechung innerhalb der Melodiestructur definiert, die mit einem Intensitätsabfall von $\geq 10\text{dB}$ einhergingen. Einfach- bzw. Mehrfachsegmentierungen (eine bzw. mehr als eine Pause) von Melodiebögen wurden hier in der Variable, „SEG“ zusammengefasst. Segmentierte Strukturen bilden hier eine sehr heterogene Gruppe, da sowohl die Anzahl der Pausen als auch deren Dauer variabel waren. Eine spezifische Differenzierung der melodisch-prosodischen Eigenschaften dieser Vokanten ist daher im Rahmen dieser Arbeit nicht vorgenommen worden.

Zusätzlich zu der beschriebenen Bogen- und Segmentierungsstruktur können Melodien (außer Kurzlaute) auch von Ein- und/oder Ausschwingvorgänge (sogenannte An- oder Endlaute) begleitet sein. Hierbei handelt es sich um kurze Abschnitte ($< 150\text{ms}$), die vor- oder nach dem ersten bzw. letzten Melodiehauptbogen erkennbar sind (für eine detaillierte Beschreibung siehe (Mampe-Keller 2013)). Das Auftreten solcher Phänomene wurde im CDAP-System dokumentiert und die Ein- und Ausschwingvorgänge separat vermessen. In der Melodiebogenanalyse bezüglich melodischer Eigenschaften wurden diese nicht einbezogen, da die unbeeinflusste Bogenzeitstruktur von Interesse war.

3.10 Vermessung der prosodie-relevanten Melodieeigenschaften

Die Vermessung der temporalen Vokant-Eigenschaften erfolgte in CDAP interaktiv mittels manueller Cursorsetzung und berücksichtigte die zuvor ermittelte Melodiestructur. In diesem Analyseschritt wurden temporale Eigenschaften aller Vokanten anhand der Melodie vermessen und automatisch in Exceltabellen gespeichert.

Im Rahmen der melodisch-prosodischen Analysen, die auch „Rhythmusanalysen“ in doppelböigen Melodien beinhaltet, wurden die in Tabelle 7 aufgeführten Messgrößen verwendet.

Tabelle 7: Prosodierelevante Messgrößen

Melodielänge (Lges)	Dauer der Melodie, unabhängig von der Bogenanzahl und ohne Berücksichtigung von Ein- und Ausschwingvorgängen
Bogenlänge LB	Vermessung der einzelnen Hauptbogenlängen (ohne Ein- und Ausschwingvorgänge)
Bogenalpha α und Bogenformen	Bogenalpha α beschreibt das zeitlich relative Auftreten des Frequenzmaximums eines Bogens im Verhältnis zur Gesamtbogenlänge (Mampe et al. 2009). Das Bogenalpha kann dabei einen Wert zwischen 0 und maximal 1 annehmen, da die Dauer der Vokanten vor der Berechnung normiert wird. Es werden 3 Bogenformen unterschieden: a) Fallend $\alpha < 0,4$ b) Symmetrisch $0,4 \leq \alpha \leq 0,6$ c) Steigend $\alpha > 0,6$
▼ „Rhythmusanalyse“ 2B-Strukturen: ▼	
LB1/Lges und Rhythmusschemata: Trochäus/ Jambus/ Levelstress	LB1/LBges stellt den Quotienten aus der Bogenlänge des ersten Bogens und der Gesamt-Melodielänge der zweiböigen Vokanten dar. Mithilfe dieses Quotienten kann das Betonungsverhalten des Säuglings innerhalb der Vokanten gemessen und je nach Wert einem entsprechenden rhythmischen Muster zugeordnet werden. Ein Wert $\leq 0,45$ impliziert eine kürzere Länge des ersten Bogens im Verhältnis zum zweiten und entspricht damit dem jambischen Muster späterer zweisilbiger Wörter, „kurz – lang“, bei dem die Betonung auf der ersten Silbe liegt. Ein Wert $\geq 0,55$ impliziert im Umkehrschluss ein trochäisches Rhythmik-Schema – „Lang – Kurz“. Levelstress beschreibt einen Doppelbogen ohne markantem Betonungsverhalten und beinhaltet Werte größer 0,45 bzw. kleiner 0,55.
Bogenkombimuster	Kombinationsmöglichkeiten und -häufigkeiten der drei Bogenformen in Doppelbögen

3.11 Statistik

Im letzten Schritt der Datenauswertung folgte die statistische Analyse mithilfe von „Microsoft Excel“ und „IBM SPSS Statistics“. Die zugrundeliegenden Daten lieferten hierbei die in CDAP erfolgten Vermessungen und Datenauswertungen.

Über eine Routine wurden die Messwerte aus CDAP anhand der manuellen Cursorsetzung, sowie der errechneten Frequenzdaten aus PRAAT inklusive aller weiteren vermerkten Lautcharakteristika in eine Exceltabelle exportiert. Anhand dieser konnte nun mittels SPSS eine statistische Datenanalyse vorgenommen werden, die im Ergebnisteil berichtet wird.

Die Darstellungen der Befunde erfolgten mittels Streudiagrammen, Histogrammen und Boxplots und wurden jeweils mit Tabellen zur deskriptiven Statistik ergänzt.

Das Streudiagramm fand im Rahmen der Formant-Auswertungen Anwendung. Es beinhaltet das in der Literatur beschriebene Vokaldreieck (Masapollo et al. 2016) und wird mit den Formantdaten (F1 und F2) unserer Vokanten ergänzt. Der genaue Hintergrund dazu ist ausführlich in Kap. 3.8 beschrieben.

Eine weitere Untersuchungsvariable ist die Melodiestructur. Unterschieden wurde hierbei zwischen Kurzlauten, Einzel- und Mehrfachbögen sowie segmentierten Strukturen (vgl. Kap. 3.9). Um den Grad der Komplexität der Melodiestructuren im Gruppenvergleich durch eine einfache Maßzahl auszudrücken, ist der Melodiekomplexitätsindex (MCI) geeignet. Er beschreibt den relativen Anteil komplexer Formen (Mehrfachbögen und segmentierte Bogenstrukturen) am gesamten Lautrepertoire (Wermke et al. 2007).

$$MCI = \frac{2B + MB + SEG}{KL + 1B + 2B + MB + SEG}$$

MCI = Melodiekomplexitätsindex
KL = Kurzlaut
1B/2B/MB = Einer-/Doppel-/Mehrfachbögen
SEG = Segmentierungen

Neben einer tabellarisch aufgeführten deskriptiven Statistik dienen Balkendiagramme zur Visualisierung der Häufigkeiten der einzelnen Strukturklassen.

Im Rahmen der melodisch-prosodischen Vergleiche dienen im Folgenden in erster Linie Balkendiagramme, die die Verteilung der skalierten Melodielängen veranschaulichen. Diese werden im Sinne einer visuell nachvollziehbareren Perzentilenverteilung in Form von Boxplots teilweise ergänzt. Differenziert werden der Median, sowie die obere und untere Quartile. Der Median entspricht dabei der 50. Perzentile, die obere und untere Quartile der 75. bzw. der 25. Perzentile.

Die Perzentile beschreibt dabei den relativen Anteil der gemessenen Größen, die sich unterhalb des angenommenen Wertes befinden. Für den Median gilt, dass sein ermittelter Wert 50% der Verteilung übertrifft. Er entspricht der 50. Perzentile. Gegenüber einer Mittelwert-Betrachtung richtet sich die Perzentile nach der relativen Verteilung einer untersuchten Variablen.

Boxplots ermöglichen zusätzlich eine grafische Darstellung von potenziellen Ausreißern und Extremwerten. Abbildung 5 veranschaulicht eine beispielhafte Skizze eines Boxplots.

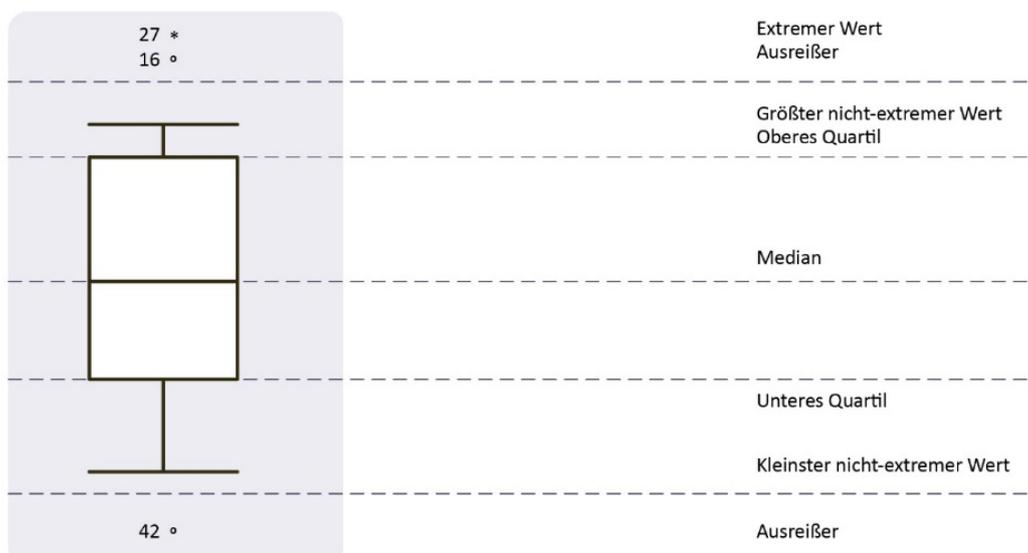


Abbildung 5: Schematische Darstellung eines Boxplots

Quelle: Newcombe (1992)

Signifikanzprüfung

Zur Beurteilung statistischer Signifikanz der Forschungsergebnisse wurde das lineare gemischte Modell der Statistiksoftware SPSS angewandt und nach Beratung durch die Statistikerin Frau Dr. Daniela Keller vom Autor durchgeführt.

Überprüft wurden die Ergebnisse hinsichtlich eines potenziell signifikanten Alters- oder Gruppeneffekts. Das Alter der Probanden wurde dabei auf 0 normiert und mit der Variablen „Alterd0“ gekennzeichnet. Für die Probandengruppen (= PG) erfolgte eine binäre Umcodierung. PG = 1 entspricht im Rahmen aller statistischen Signifikanzprüfungen der Kontrollgruppe, PG = 2 den Patientengruppe. Nach Voruntersuchungen lieferte der Interaktionsterm der berichteten Faktoren in keinem Fall ein Signifikantes Ergebnis, weshalb dieser in den folgenden Modellen nicht berücksichtigt wurde und die Haupteffekte – Alter und Gruppe - interpretierbar waren.

Auswertungen der Melodiekomplexität erfolgten ebenfalls mit Hilfe verallgemeinerter gemischter Modelle auf binärer Ebene; „0“ für Kurzlaute und Einerbögen, „1“ für Doppel- und Mehrfachbögen und Segmentierte Vokanten (vgl. MCI).

Auch die statistischen Untersuchungen der metrischen Beobachtungen, explizit der Melodielängenuntersuchungen der Vokanten, unabhängig von ihrer Melodiestructur, erfolgten mittels linearem gemischtem Modell. Aufgrund einer zu geringen Datenlage wurden die Melodielängen der doppelböigen Vokanten rein deskriptiv untersucht (Kap. 4.3.3).

Voraussetzungen für das gemischte Modell sind eine angemessene Normalverteilung der zu untersuchenden Größe, sowie eine entsprechende Homoskedastizität. Zunächst wurden also alle insgesamt erfassten Melodielängen aller Vokanten unabhängig ihrer Melodiestructur auf Normalverteilung mittels Q-Q-Diagramm überprüft. Nach Transformation zeigte sich eine hinreichende Normalverteilung bezüglich der In-logarithmierten Melodielängen sowie der Residuen (vgl. Abbildung 6).

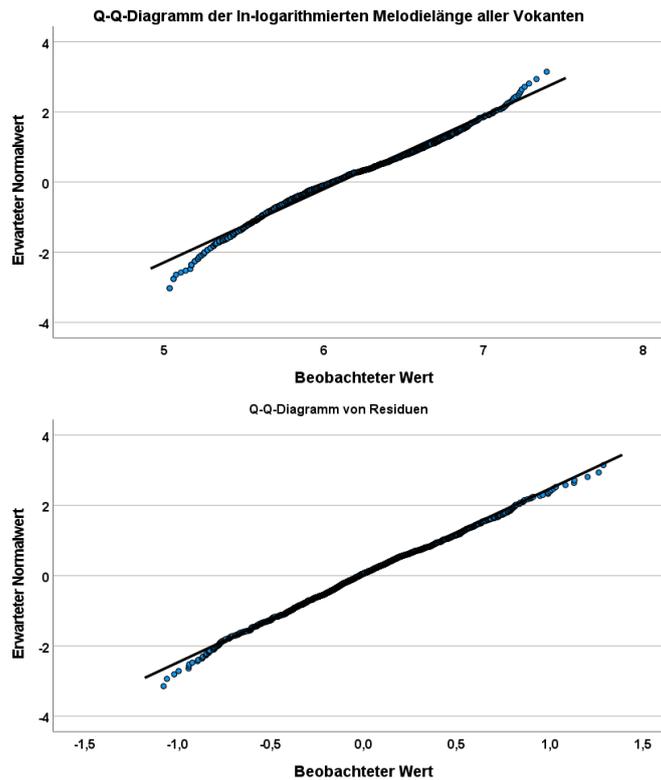


Abbildung 6: Q-Q-Diagramm der In-logarithmierten Melodielängen aller Vokanten (oben) und der Residuen der Melodielängen (unten)

Zur ergänzenden Überprüfung der Homoskedastizität unserer Daten wurden die Prädiktoren den Residuen des gemischten Modells (ohne Interaktionsterm) im Streudiagramm gegenübergestellt (Abbildung 7).

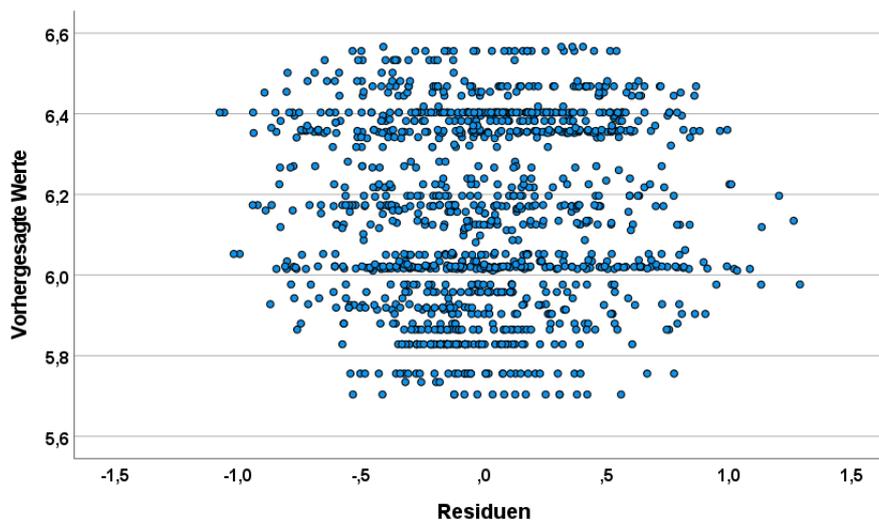


Abbildung 7: Prädiktor-Residuen-Streudiagramm des gemischten Modells (ohne Interaktionsterm) der Melodielängen als abhängige Variable unter Berücksichtigung aller Vokanten

In Abbildung 7 zeigt sich eine unregelmäßig angeordnete Punktwolke, es sind keine erkennbaren Tendenzen der Prädiktor-Residuen-Beziehungen zu beobachten. Somit konnte auch von einer hinreichenden Homoskedastizität ausgegangen werden (vgl. Abbildung 7).

Die Kriterien für die Anwendung eines linearen gemischten Modells wurden demnach hinreichend erfüllt. Zur weiteren Absicherung der Ergebnisse wurde dennoch innerhalb der metrischen statistischen Analysen zusätzlich ein Bootstrapping durchgeführt. Mit einer beliebig hohen Stichprobenwahl konnte hierbei der, wenn auch geringe, Einfluss der Abweichungen in der Normalverteilung bereinigt werden. So bestätigten die Messwiederholungen im Rahmen des Bootstrappings, neben dem linearen Modell, eine zusätzliche, methodisch unabhängige Signifikanz. Sofern nicht anders angegeben, beruhen die Bootstrap-Ergebnisse in allen folgenden Untersuchungen auf 5000 Bootstrap-Stichproben.

Auch hierbei wurde eine mehrfaktorielle Überprüfung hinsichtlich des Alters- und Gruppeneffekts durchgeführt.

Ergänzend zur tabellarischen Auswertung ist das Akaike-Informationskriterium (AIC) sowie die Varianz des zufälligen Effekts angegeben. Beim AIC handelt es sich um eine einheitslose Variable. Sie ist das Gütekriterium bei der Bewertung statistischer Modelle.

Der p-Wert ist auch einheitslos und gilt als Richtwert für die Signifikanz einer Beobachtung. Fällt dieser unter den Wert von 0,05, galt der untersuchte Zusammenhang hier als statistisch signifikant.

4 Ergebnisse

Das Projekt, in das die Arbeit eingebettet ist, beinhaltet den Vergleich von Komfortlauteigenschaften zwischen sensorineural hörgestörten und gesunden Säuglingen. Der Fokus der vorliegenden Arbeit lag auf der Gegenüberstellung der zwei kontraststärksten Gruppen der Kohortenstudie, der hochgradig sensorineural hörgestörten Säuglinge (HI-Gruppe) und der normalhörenden Säuglinge (KG-Gruppe) (Kap. 3.4). Für den ersten Vergleich dieser Gruppen wurde hier der Vokalisationstyp „Vokant“ analysiert (vgl. Kap. 3.7.2)

Der Ergebnisteil ist wie folgt aufgebaut:

Zuerst erfolgt die Darstellung der Ergebnisse der Formantanalyse (vgl. 3.8.2). Dabei wurden von audiovisuell identifizierten Vokanten Formantwerte berechnet und die gemittelten Werte für F1 und F2 im Vokaldreieck eingezeichnet.

In einem weiteren Kapitel wird ein Überblick über die verschiedenen beobachteten Melodiestructuren innerhalb der Vokanten gegeben und das Ergebnis der Gegenüberstellung der Melodiekomplexitäten von Vokanten beider Gruppen berichtet.

Der letzte Teil berichtet temporale Messgrößen und Formeigenschaften der am häufigsten beobachteten Melodiestructurkategorien, den ein- und zweibögigen Vokanten. Dieser Teil beinhaltet auch die Ergebnisse der Analyse einfacher Rhythmus-Charakteristika in doppelböigen Vokanten.

4.1 Formantanalyse

Die, wie in Kapitel 3.8 beschrieben, ermittelten Resonanzfrequenzen (Formanten F1, F2) in Vokanten sind in Abbildung 8 in Form eines Vokalraumdiagramms dargestellt. Die für ältere Säuglinge typischen Frequenzen der Eck-Vokale (/i/, /a/ und /u/) sind zusätzlich eingezeichnet, da es für Vokanten bisher keine Referenzwerte in der Fachliteratur gibt. Die Gesamtanzahl der in Abbildung 8 und Abbildung 9 dargestellten Formantdaten beträgt N= 683.

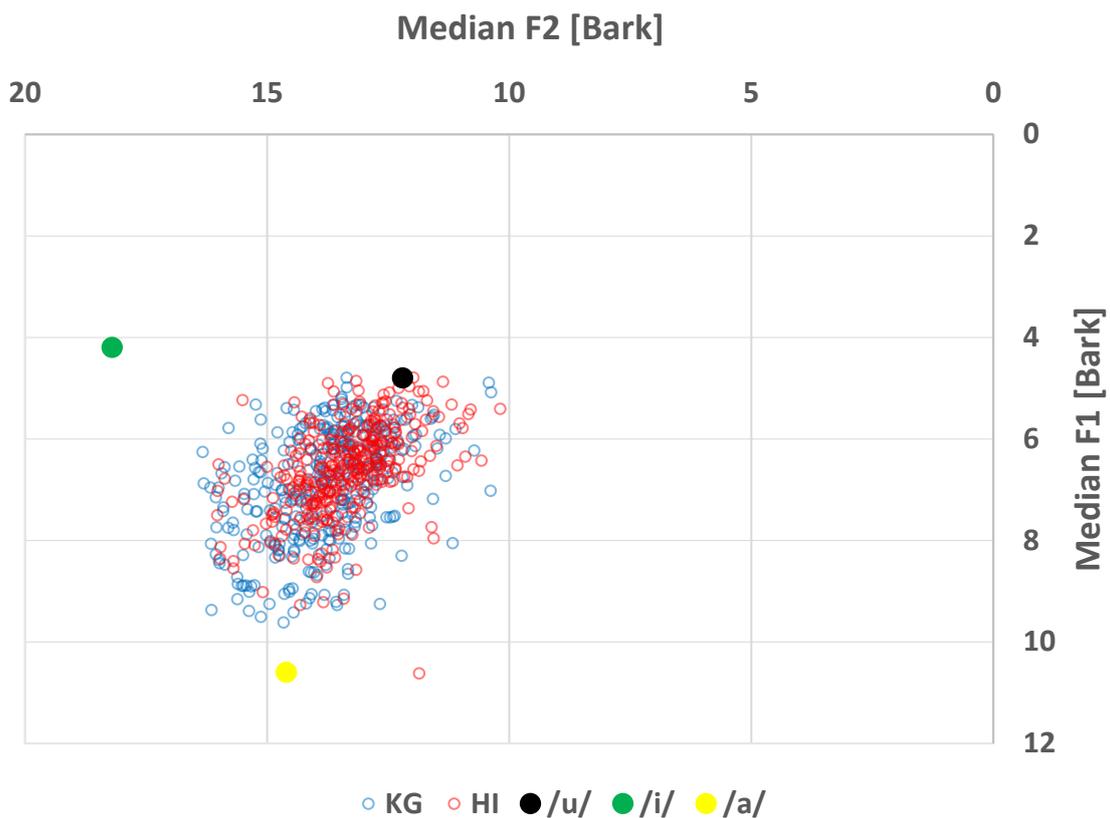


Abbildung 8: Verteilung der Formantfrequenzen F1 und F2 der Vokanten im Vokalraumdiagramm

Die Werte für den ersten (F1) und zweiten Formanten (F2) sind psychoakustisch korrekt in einer Bark-Skala, wie dargestellt, eingezeichnet. Die Werte der Kontrollgruppe sind durch blaue, die der hörgestörten Säuglinge durch rote Kreise markiert. Die Referenzwerte für das Vokaldreieck bilden die Formantwerte der Vokale /i/, /u/ und /a/ für Säuglinge im Alter von 4-6 Monaten, berichtet in Masapollo et al. (2016).

In Abbildung 8 sind verschiedene Befunde erkennbar:

- (1) Die Formantwerte beider Gruppen überlagern sich.
- (2) Es gibt ein Cluster dessen Zentrum bei den Werten von ca. 6 Bark (F1) und ca. 13 Bark (F2) liegt.
- (3) Die KG-Gruppenwerte zeigten eine größere Tendenz der Werte in Richtung des Vokals /a/, also eines vollen Vokals.

In Abbildung 9 sind die gemittelten Formantwerte F1 und F2 in Form von Boxplots beider Gruppen gegenübergestellt.

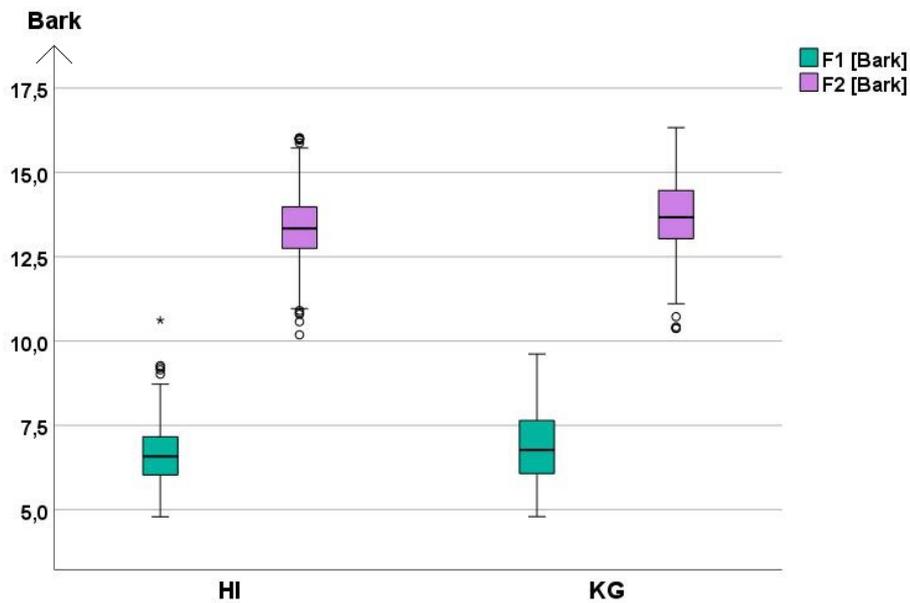


Abbildung 9: Boxplots der Formantfrequenzen F1 und F2 der Vokanten der Gruppe der hochgradig sensorineural gestörten Säuglinge HI (N=331 Vokanten) und der Gruppe der normalhörenden Säuglinge (KG) (N=352 Vokanten)

Tabelle 8 liefert eine deskriptive Statistik der über alle Vokanten gemittelten Grundfrequenzen (f_0) und der ersten beiden Resonanzfrequenzen (F1 und F2) beider Gruppen.

Tabelle 8: Deskriptive Statistik der gemittelten Grund- (f_0) und Resonanzfrequenzen (F1 und F2) der Vokanten

Gruppe	Messgröße	f_0 [Hz]	F1 [Hz]	F2 [Hz]
HI (N= 331)	N	331	331	331
	M	344,45	734,80	2114,72
	Md	350,86	724,74	2075,03
	SD	36,43	121,89	340,00
	Min	221,66	506,44	1290,14
	Max	399,84	1375,93	3226,86
KG (N= 352)	N	352	352	352
	M	325,98	775,71	2260,07
	Md	319,69	750,16	2186,24
	SD	35,26	148,07	409,81
	Min	237,89	506,88	1326,72
	Max	399,49	1184,83	3401,72

Die statistische Untersuchung der hier ermittelten Formantfrequenzen erfolgte mittels verallgemeinertem gemischtem Modell (vgl. Kap. 3.11). Die Ergebnisse dieser sind in Tabelle 9 für die gemittelten Werte der ersten und zweiten Formanten dargestellt.

Untersucht wurde auf signifikante Unterschiede zwischen den beiden Probandengruppen.

Als Gütekriterium für das Modell wurde der Akaikewert (AIC) verwendet (vgl. Kap. 3.11). Er beträgt 8611,46 bei der Untersuchung der ersten Formanten bzw. 9983,78 bei der Untersuchung der zweiten.

Tabelle 9: Verallgemeinertes gemischtes Modell zur Untersuchung des Gruppeneffektes auf die pro Vokant gemittelten Werte des ersten (F1) und zweiten Formanten (F2)

	Parameter	AIC	Schätzung	Signifikanz	KI95%	
					UG	OG
F1	Konstanter Term [PG=1] (HI=2)	8611,46	753,8	<0,001	716,63	790,96
			-16,70	0,552	-74,58	41,18
F2	Konstanter Term [PG=1] (HI=2)	9983,78	2226,74	<0,001	2129,53	2323,95
			-109,61	0,144	-259,53	40,32

Die p-Werte liegen mit 0,552 für die ersten Formanten bzw. 0,144 für die zweiten Formanten weit außerhalb des Signifikanzniveaus von <0,05. Signifikante Gruppenunterschiede bei den gemittelten ersten und zweiten Formantfrequenzen bestehen also nicht.

4.2 Melodiestrukturanalyse

Die Ergebnisse der Melodiestrukturanalyse (vgl. Kap. 3.9) für die Vokanten der beiden Gruppen (HI und KG) über den gesamten Untersuchungszeitraum (N= 1236) sind in Abbildung 10 in Form von Säulendiagrammen der beobachteten Häufigkeiten dargestellt. Einböigige, zweiböigige und segmentierte Melodiestrukturen wurden hierbei jeweils zusammengefasst und gegenübergestellt. Mehrböigige unsegmentierte Melodien (> 2B) wurden nur in acht Fällen beobachtet und waren somit eine Ausnahme. Tabelle 10 berichtet diese Einzelfälle der entsprechenden Säuglinge („>2B“); in Abbildung 10 sind diese nicht aufgeführt.

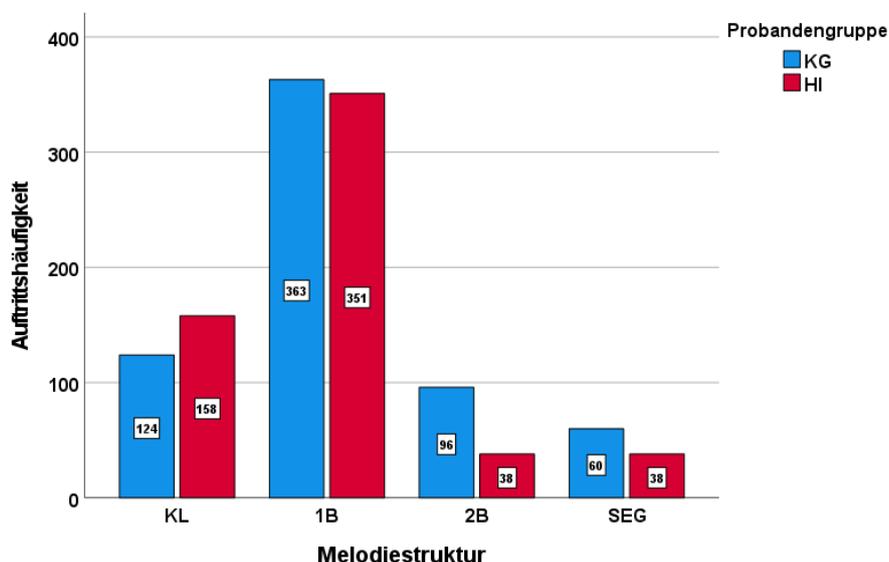


Abbildung 10: Melodiestrukturverteilung der Vokanten der Gruppe HI und KG

Die absoluten Auftrittshäufigkeiten über den Untersuchungszeitraum einfacher Melodiestrukturen (Einzelbögen (1B) und Kurzlaute (KL)) und komplexer Melodien sind vergleichend zwischen der Kontrollgruppe (KG) und der hörgestörten Gruppe (HI) dargestellt. Bei den komplexen Formen sind die zweiböigigen Strukturen (2B) differenziert von den ein- und mehrfach segmentierten Strukturen (SEG) berichtet.

Tabelle 10: Beobachtete mehrböigige Vokanten über den gesamten Untersuchungszeitraum

Gruppe	Säugling	Auftrittshäufigkeit „>2B“
HI	BA04	1
	BB02	2
KG	AA03	2
	AB01	2
	AB02	1

In Tabelle 11 sind zusätzlich zu den absoluten auch die relativen Häufigkeiten der beobachteten Strukturen berichtet. Um die Komplexität der Melodiestructur durch eine einfache Maßzahl zu beschreiben, wurde auch der Komplexitätsgrad für jede Gruppe angegeben. Dieser entspricht dem Anteil komplexer (2B + SEG) gegenüber einböiger Strukturen (1B+KL).

Tabelle 11: Absolute und relative Häufigkeit beobachteter Strukturkategorien

Gruppe	Melodie- struktur	Absolute Häufigkeit	Relative Häufigkeit	Melodie- komplexität
HI	KL	158	26,87%	13,44%
	1B	351	59,69%	
	2B	38	6,46%	
	SEG	38	6,46%	
	MB	3	0,51%	
KG	KL	124	19,14%	24,85%
	1B	363	56,02%	
	2B	96	14,81%	
	SEG	60	9,26%	
	MB	5	0,77%	

Aus Tabelle 11 geht hervor, dass die Melodiekomplexität der Vokanten in der KG-Gruppe mit 25% annähernd doppelt so hoch ist, wie die der HI-Gruppe mit 13%. Dies ist in erster Linie auf den Anteil der Vokanten mit doppelböiger Melodie (2B) zurückzuführen, deren Anteil am berichteten Repertoire doppelt so hoch war wie bei der HI-Gruppe. Der Anteil segmentierter Strukturen ist unter den normalhörenden Säuglingen nur leicht erhöht im Vergleich zu den hörbehinderten (9,26% gegenüber 6,46%).

Die Zahlen aus Tabelle 11 zeigen, dass der Anteil der komplexen Strukturen vor allem bei zweiböigen Melodien zwischen den Gruppen variiert. Vor diesem Hintergrund wurden die melodisch-prosodischen Eigenschaften differenziert nur für die am häufigsten beobachteten Melodiestructuren in Vokanten - die einfachen Strukturen (1B und KL) und die Melodien mit Doppelbögen (2B) - in den nachfolgenden Kapiteln berichtet.

Zur statistischen Analyse der Beobachtungen wurde ein verallgemeinertes gemischtes Modell angewendet. Tabelle 12 zeigt die Ergebnisse der statistischen Prüfung mithilfe des gemischten Modells bezüglich der Melodiekomplexität und deren Verteilung in Abhängigkeit vom Alter und der Probandengruppe (vgl. Kap. 3.11).

Die Analyse erfolgte mit Hilfe einer binären Variablen (0/1) für die Melodiekomplexität. Die Variable „Alterd0“ stellt das Alter der Probanden dar, wobei das jüngste der Studie vorliegende Alter (= 60 Tage) als Ursprung gewählt wurde (= Tag 0). Somit ist die Variable Alterd0 auf 0 normiert und entspricht der Differenz des Alters des Probanden am Aufnahmetag und den 60 Tagen (vgl. Kap. 3.11).

Insgesamt gingen 1236 Vokanten in diese Untersuchung mit ein, davon waren 996 nicht komplex. Der Akaikewert (AIC) (vgl. Kap. 3.11) beträgt 5931,639. Der zufällige Effekt hatte eine Varianz von 0,250 mit einem p-Wert von 0,122.

Tabelle 12: Verallgemeinertes gemischtes Modell zur Untersuchung des Gruppeneffektes auf die binäre Größe für die Melodiekomplexität

Modellterm	T	Signifikanz	OR	Konfidenzintervall 95%	
				Untergrenze	Obergrenze
Konst. Term	-5,779	<0,001	0,070	0,028	0,175
Alterd0	3,446	0,001	1,015	1,006	1,024
PG=1 (HI = 2)	2,047	0,044	2,791	1,028	7,577
Alterd0*[PG=1]	-1,808	0,073	0,991	0,981	1,001

Mit einem p-Wert von 0,001 hat das Alter einen hoch signifikanten Einfluss auf die Melodiekomplexität. Die „Effektstärke“ (OR / Odds Ratio) beträgt dabei 1,015.

Der Gruppeneffekt auf die Melodiekomplexität belief sich knapp unterhalb des Signifikanzniveaus von 0,05 ($p=0,044$) und bestätigt damit auch hier einen statistisch signifikanten Zusammenhang.

Eine signifikante Interaktion zwischen Alter und Probandengruppe bezüglich der Melodiekomplexität wurde mit $p = 0,073$ nicht festgestellt.

4.3 Melodisch-prosodische Eigenschaften von Vokanten

Zentraler Untersuchungsgegenstand der vorliegenden Arbeit waren die melodisch-prosodischen Eigenschaften der Vokanten. Im Folgenden werden die Ergebnisse der Analyse der Melodielänge aller Vokanten berichtet.

Wie aus Abbildung 10 in Kapitel 0 erkennbar, wurden in Vokanten über den Untersuchungszeitraum in beiden Gruppen vor allem Kurzlaute sowie ein- (1B) und doppelböigige (2B) Melodien beobachtet. Bei der Analyse der Melodielängen und anderer temporaler Größen wurden hier daher vor allem diese Vokanten berücksichtigt. 2B-Melodien wurden aufgrund ihrer spezifischen Analogie zu Doppelsilben in den späteren vokalen Entwicklungsphasen auch in einer Art linguistischen „Rhythmusanalyse“ ausgewertet (4.4).

4.3.1 Melodielängen aller Vokanten

Die zeitliche Dauer der Vokanten, also die Melodielänge, wurde ohne Berücksichtigung der Ein- und Ausschwingvorgänge (vgl. Kap. 3.9) berechnet. In die Auswertung kamen 1209 (HI Gruppe: 578; KG-Gruppe: 631) der insgesamt 1236 Vokanten. Bei 27 Vokanten konnte die Melodielänge aufgrund phonatorischer Rauschbanden am Anfang oder am Ende des Vokanten nicht verlässlich ermittelt werden.

Die vergleichende Analyse für diese Variable ist in Form von Boxplots in Abbildung 11 sowie ergänzend mit Histogrammen beider Gruppen in Abbildung 12 und Abbildung 13 dargestellt. Die zugehörige deskriptive Statistik findet sich in Tabelle 13. Die Darstellungen in Abbildung 11 bzw. Abbildung 12 und Abbildung 13 sind in gewisser Weise redundant, sie machen aber die beobachteten Messungen insgesamt vergleichbarer.

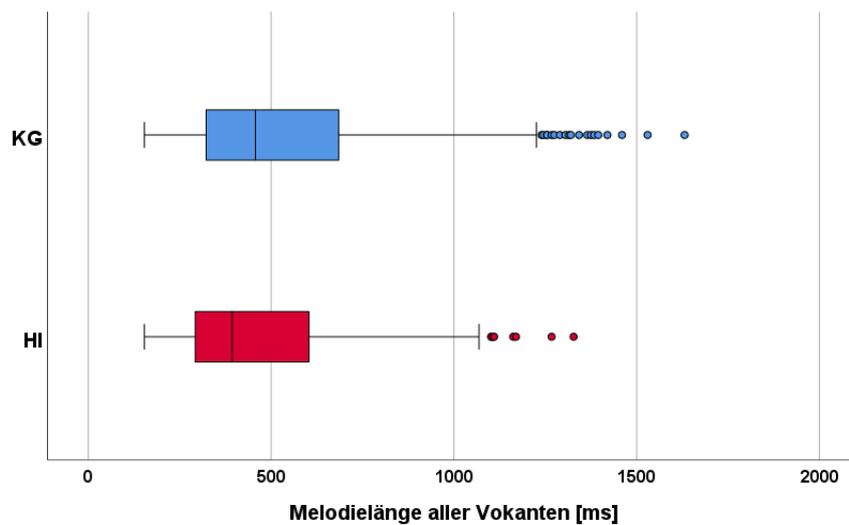


Abbildung 11: Boxplots für die Melodielänge der Vokanten der Kontrollgruppe (KG) und der Gruppe der hörgestörten Säuglinge (HI)

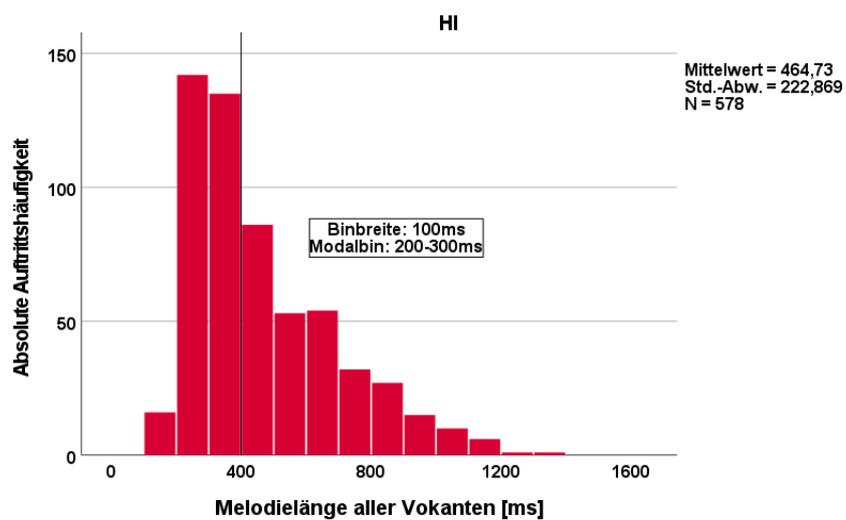


Abbildung 12: Längenverteilung der Vokanten der Gruppe HI unabhängig von der Melodiestructur

Die senkrechte Orientierungslinie ($x=400\text{ms}$) markiert den in der Literatur angegebenen häufigsten Maximalwert für vokalische Äußerungen im Untersuchungsalter (Bauer und Kent 1987; Kent und Murray 1982).

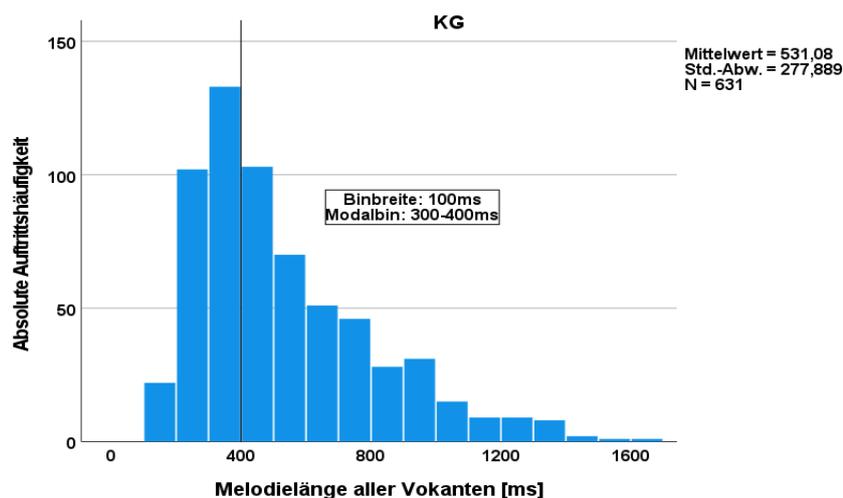


Abbildung 13: Längenverteilung der Vokanten der Gruppe KG unabhängig von der Melodiestructur

Die senkrechte Orientierungslinie ($x=400\text{ms}$) markiert den in der Literatur angegebenen häufigsten Maximalwert für vokalische Äußerungen im Untersuchungsalter (Bauer und Kent 1987; Kent und Murray 1982).

Tabelle 13: Melodielänge aller Vokanten ohne Berücksichtigung der Melodiestructur

Gruppe	N	Melodielänge [ms]					CI95%	
		Min	Max	M	SD	Md	UG	OG
HI	578	154	1328	465	223	392	447	483
KG	631	154	1631	531	278	458	509	553

N= Anzahl; Min = Minimum; Max = Maximum; M = Mittelwert; SD = Standardabweichung; Md = Median; CI = Konfidenzintervall; UG = Untergrenze; OG = Obergrenze

Aus den Diagrammen und den berichteten Werten geht hervor, dass die Melodielängen der Säuglinge der Kontrollgruppe im Median um ca. 66ms länger sind als die der Säuglinge der Gruppe HI (HI: 392ms; KG: 458ms).

Die vermessenen Minimalwerte der untersuchten Vokanten beider Gruppen waren mit 154ms identisch, wobei die im Rahmen der Analyse festgelegte Mindestlänge von 150ms (vgl. Kap. 3.9) einen methodisch bedingten Grenzwert lieferte.

Da es bisher in der Literatur keine Angabe von Vokantenlängen im Untersuchungszeitraum gibt, wurde in Tabelle 14 neben der deskriptiven Gruppenstatistik auch die individuellen Messwerte berichtet, um sie nachfolgenden Studien zur Verfügung zu stellen. Sie werden auch vom Autor publiziert.

Tabelle 14: Deskriptive Statistik der individuellen Messwerte für die Melodielängen

Gruppe	Säugling	N	Melodielänge [ms]						
			Min	Max	M	SD	Md	CI95%	
								UG	OG
HI	BA01	18	240	1069	517	285	411	375	659
	BA02	31	210	521	366	67	364	341	390
	BA03	37	259	1043	587	178	589	528	646
	BA04	112	154	1328	625	233	608	581	668
	BB01	38	285	1108	691	227	653	618	765
	BB02	107	176	1110	436	210	370	396	477
	BB03	177	165	904	335	123	307	317	354
	BB04	58	203	885	416	169	364	372	461
KG	AA01	17	170	787	366	174	315	276	455
	AA02	3	240	300	274	31	281	197	350
	AA03	3	619	934	820	175	908	386	1254
	AA04	2	319	394	356	53	356	120	832
	AA05	10	157	776	385	217	333	229	540
	AA06	40	244	1050	589	219	578	519	659
	AA07	10	255	965	533	229	544	369	697
	AA08	1							
	AA09	7	180	1395	449	429	248	53	846
	AA10	4	188	259	225	31	227	176	274
	AA11	9	251	758	439	181	364	300	578
	AA12	14	188	581	419	118	396	351	487
	AA13	7	240	514	325	108	266	224	425
	AA14	17	154	878	444	224	454	329	560
	AA15	7	229	926	415	236	319	197	633
AB01	190	175	1631	465	237	390	431	499	
AB02	160	160	1460	561	292	475	516	607	
AB03	130	210	1530	673	309	585	620	726	

N= Anzahl; Min = Minimum; Max = Maximum; M = Mittelwert; SD = Standardabweichung; Md = Median; CI = Konfidenzintervall; UG = Untergrenze; OG = Obergrenze

Die statistische Untersuchung der Melodielängen erfolgte mittels gemischtem Modell (vgl. Kap. 3.11). Tabelle 15 zeigt die untersuchten Parameter der ln-logarithmierten Melodielängen aller Vokanten in Abhängigkeit vom a) Alter (Alterd0) und der b) Probandengruppe (PG) hinsichtlich Signifikanz nach Bootstrap-Analyse (vgl. 3.11).

Nach 5000 zufälligen Bootstrap-Stichproben betrug der AIC hierbei 1416,497. Der Schätzer der Varianz des zufälligen Effekts lag bei 0,047.

Tabelle 15: Bootstrap-Analyse der ln-Logarithmierten Melodielängen aller Vokanten

Parameter	Schätzung	Verzerrung	Signifikanz	KI95%	
				UG	OG
Konstanter Term	5,982	<0,001	<0,001	5,920	6,046
[PG=1] (HI=2)	-0,107	-0,001	0,037	-0,181	-0,036
Alterd0	0,002	<0,001	<0,001	0,002	0,003

Mit einem p-Wert von 0,037 zeigt sich bezüglich der Melodielängen unter Berücksichtigung aller Lautkategorien ein signifikanter Gruppenunterschied. Mit $p < 0,001$ ist auch ein hoch signifikanter positiver Alterseffekt zu beobachten. Die Schätzungen ergeben nach Rückrechnung einen Faktor von 0,898 ($e^{-0,107}$) für den Gruppen- bzw. 1,002 ($e^{0,002}$) für den Alterseffekt. Die Potenzen der durch die Bootstrapping-Methode bedingt mögliche Verzerrung betragen dabei -0,001 bzw. unter 0,001 und sind demnach äußerst gering einzustufen.

Die Signifikanz des Interaktionsterms lag nach Voruntersuchungen mit $p=0,11$ außerhalb der Signifikanz-Grenze. Es zeigen sich demnach keine signifikanten Gruppenunterschiede bezüglich der angenommenen altersabhängigen Melodielängenzunahme.

4.3.2 Melodielänge in einfach strukturierten Vokanten (1B)

Die folgenden Abbildungen 15-17 zeigen Melodielängen-Verteilungen der einbögigen Vokanten (inklusive Kurzlaute) der Gruppen KG und HI. Zu beachten ist auch hierbei, dass Vokanten erst ab einer Länge von mindestens 150ms als solche klassifiziert wurden (vgl. Kap. 3.9). Kurzlaute (Länge 150-300ms) zeigen neben einfachen (einbögigen) Melodiestructuren teilweise auch stark modellierte Konturen. Sie wurden daher auch separat von den einbögigen Vokanten (> 300ms) betrachtet und in Abbildung 14 gesondert aufgezeigt.

Die Abbildungen 16 und 17 stellen die Ergebnisse in Form von Histogrammen für beide Gruppen gegenüber.

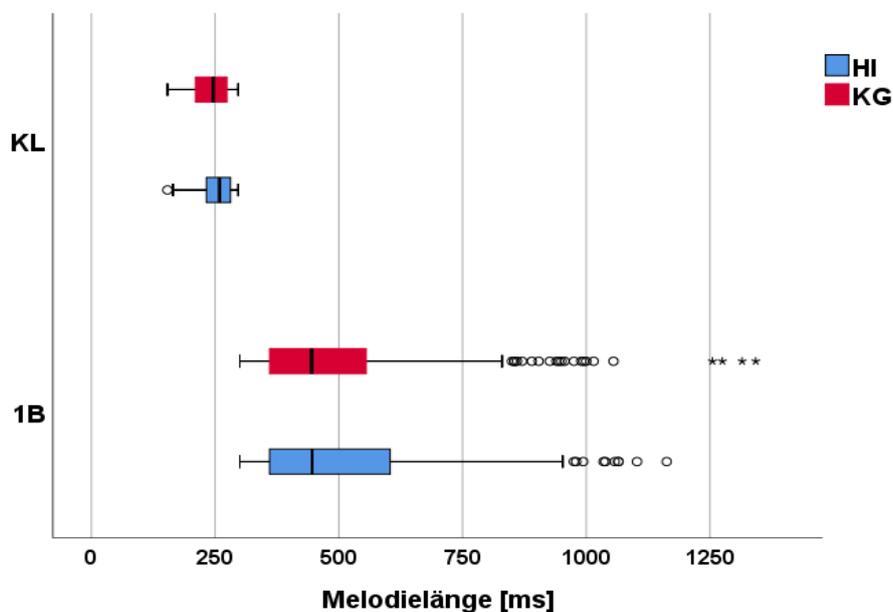


Abbildung 14: Boxplot der Melodielängen für einbögige (1B) Vokanten und Kurzlaute (KL)

Werte der HI-Gruppe (HI) werden in Form der roten Boxplots, Werte der Kontrollgruppe (KG) in blauen Boxplots markiert.

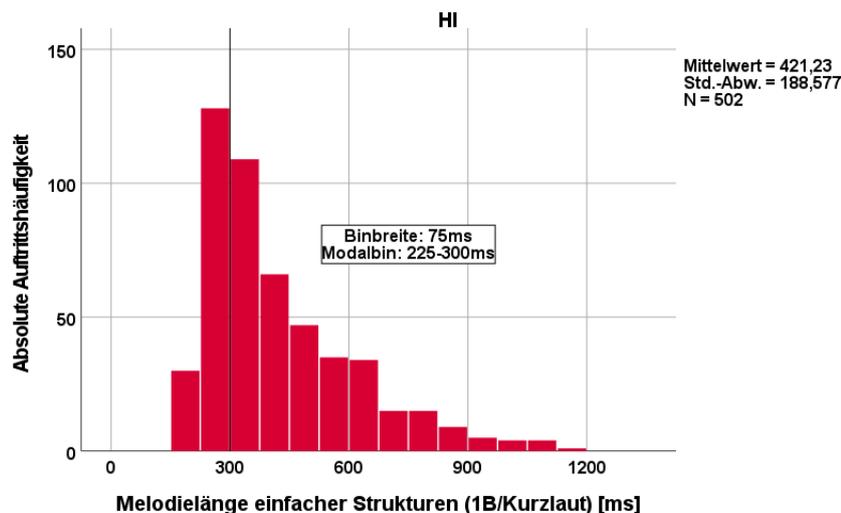


Abbildung 15: Längenverteilungen der Vokanten mit einfacher Melodiestructur der Gruppe HI

Dargestellt sind die Messwerte aller Vokanten mit einbögiger Melodie, die im Untersuchungszeitraum von Säuglingen der HI-Gruppe ausgewertet wurden. Die senkrechte Orientierungslinie ($x=300\text{ms}$) markiert den hier definierten Grenzbereich zwischen Kurzlauten bis 300ms Länge und einbögigen Melodien $\geq 300\text{ms}$ Länge.

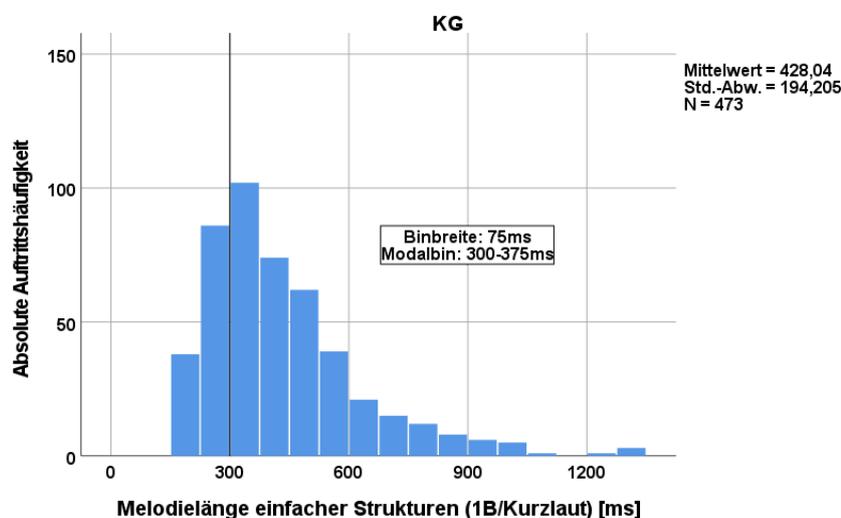


Abbildung 16: Längenverteilungen der Vokanten mit einfacher Melodiestructur der Gruppe KG

Dargestellt sind die Messwerte aller Vokanten mit einbögiger Melodie, die im Untersuchungszeitraum von Säuglingen der Kontrollgruppe ausgewertet wurden. Die senkrechte Orientierungslinie ($x=300\text{ms}$) markiert den hier definierten Grenzbereich zwischen Kurzlauten bis 300ms Länge und einbögigen Melodien $\geq 300\text{ms}$ Länge

Tabelle 16 fasst die Ergebnisse mittels deskriptiver Statistik zusammen.

Tabelle 16: Deskriptive Statistik für alle vermessenen Melodielängen einfach strukturierter Vokanten gesamt und ohne Berücksichtigung der Kurzlaute (o KL)

Gruppe	N	Melodielänge [ms]						CI95%	
		Min	Max	M	SD	Md	UG	OG	
HI	502	154	1163	421	189	364	405	434	
KG	473	154	1343	428	194	386	411	446	
HI o KL	344	300	1163	499	179	446	480	518	
KG o KL	343	300	1343	498	184	450	478	517	

N= Anzahl; Min = Minimum; Max = Maximum; M = Mittelwert; SD = Standardabweichung; Md = Median; CI = Konfidenzintervall; UG = Untergrenze; OG = Obergrenze

Man erkennt in allen Verteilungen und der Wertetabelle eine hohe Übereinstimmung zwischen den einbögigen Vokantenlängen beider Gruppen. Betrachtet man die Melodielängen der einbögigen Vokanten oberhalb des definierten Kurzlautbereichs (bis 300ms, vgl. 3.9), so unterscheiden sich die Medianwerte um lediglich 4ms (HI: 446ms; KG: 450ms), die Mittelwerte sogar nur um eine Millisekunde (HI: 499ms; KG: 498ms).

Unter Mitberücksichtigung der Kurzlautlängen unter den einbögigen Vokanten stellt sich der Median der Melodielängen der Kontrollgruppe mit 386ms gegenüber 364ms der Gruppe KG geringfügig länger dar.

Tabelle 17 stellt die Ergebnisse der abschließenden Bootstrap-Analyse eines gemischten Statistik-Modells (vgl. 3.11) der Melodielängen einbögiger Vokalisationen ohne Berücksichtigung der Kurzlaute dar.

Mit 5000 zufälligen Bootstrapanalysen der einbögigen Melodielängen betrug der AIC hierbei 752,512. Der Schätzer der Varianz des zufälligen Effekts betrug 0,042.

Tabelle 17: Bootstrap-Analyse der ln-Logarithmierten Melodielängen der Einerbögen (1B)

Parameter	Schätzung	Verzerrung	Signifikanz	KI 95%	
				Unterer Wert	Oberer Wert
Konstanter Term	5,944	0,001	<0,001	5,888	6,005
[PG=1] (HI=2)	-0,184	<0,001	<0,001	-0,248	-0,124
Alterd0	0,002	<0,001	<0,001	0,001	0,003

Der mit $p < 0,001$ hier hoch signifikante Gruppeneffekt und mit $p < 0,001$ ebenso hochsignifikante Alterseffekt ist auch bei einbögigen Vokanten ($> 150\text{ms}$) festzustellen (vgl. Tabelle 16, 4.3).

Nach Rückrechnung der ln-logarithmierten Schätzungen liefert dieser für den Alterseffekt den Faktor 1,002 ($e^{0,002}$), bzw. 0,832 ($e^{-0,184}$) für den Gruppeneffekt. Die Verzerrungen sind jeweils mit $e^{<0,001}$ auch hier sehr gering.

Nach Voruntersuchungen ließen sich mit $p = 0,329$ kein signifikantes Ergebnis für den Interaktionsterm feststellen.

4.3.3 Melodielänge der Vokanten mit doppelböigen Melodien (2B)

Insgesamt kamen die komplexen Strukturtypen in den Vokanten im Vergleich zu den einfachen selten vor (vgl. 0). Die Doppelbögen stellten hier die häufigste Form der komplexen Strukturen innerhalb der Vokanten dar. Dennoch wurden diese nicht bei jedem Säugling am Untersuchungstag beobachtet (rot markierte Fälle in Tabelle 18). Tabelle 18 berichtet die individuellen Aufnahmetage der Säuglinge, sowie den Anteil der an diesen beobachteten doppelböigen Vokanten.

Tabelle 18: Individuelle Auftrittshäufigkeit der Vokanten mit doppelböigen Melodien

Gruppe	Proband	Anzahl Aufnahmetage	Anzahl Vokanten gesamt	Anzahl Doppelbögen absolut	Anteil Doppelbögen relativ
HI	BA01	2	20	2	10%
	BA02	1	32	0	0%
	BA03	1	37	3	8%
	BA04	1	113	6	5%
	BB01	6	39	6	15%
	BB02	21	110	9	8%
	BB03	6	179	8	4%
	BB04	6	58	4	7%
KG	AA01	1	19	1	5%
	AA02	1	3	0	0%
	AA03	1	3	1	33%
	AA04	1	2	0	0%
	AA05	1	11	1	9%
	AA06	1	43	12	28%
	AA07	1	11	1	9%
	AA08	1	1	0	0%
	AA09	1	7	1	14%
	AA10	1	5	0	0%
	AA11	1	10	0	0%
	AA12	1	14	3	21%
	AA13	1	8	0	0%
	AA14	1	18	4	22%
	AA15	1	7	1	14%
AB01	16	193	21	11%	
AB02	16	160	32	20%	
AB03	10	133	18	14%	

Die berechneten Melodielängen der zweibögigen Vokanten sind in Form von Boxplots in Abbildung 17 dargestellt.

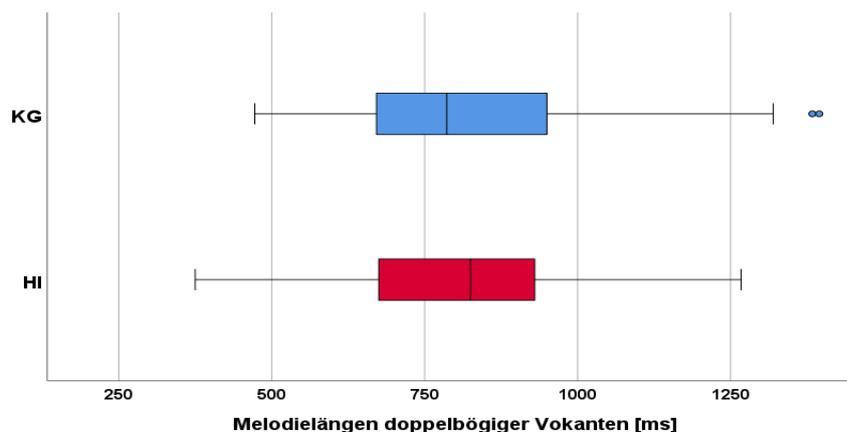


Abbildung 17: Boxplot für die Melodielängen doppelböiger Vokanten für die Kontrollgruppe (KG) und die Gruppe der hochgradig sensorineural hörgeschädigten Säuglinge (HI)

Tabelle 19 fasst die Ergebnisse der Melodielängen doppelböiger Vokanten zusammen.

Tabelle 19: Deskriptive Analyse der Melodielängen doppelböiger Vokanten

Gruppe	N	Melodielänge [ms]						
		Min	Max	M	SD	Md	CI95%	
							UG	OG
HI	38	454	1268	842	185	840	780	903
KG	96	473	1395	830	206	786	788	871

N= Anzahl; Min = Minimum; Max = Maximum; M = Mittelwert; SD = Standardabweichung; Md = Median; CI = Konfidenzintervall; UG = Untergrenze; OG = Obergrenze

Aus den Boxplots in Abbildung 17 und der Tabelle 19 geht hervor, dass die Vokanten der HI-Gruppe mit doppelböiger Melodie gegenüber der Kontrollgruppe geringfügig länger waren. Unter Berücksichtigung einer Messungenauigkeit von ± 10 ms, sind die Mittelwertbetrachtungen beider Gruppen nahezu übereinstimmend (842ms versus 830ms). Für ein multimodal logistisch aussagekräftiges Modell lag mit $N = 134$ ein zu geringer Datensatz der einzelnen Untersuchungskriterien für die 2-Böigen Vokanten vor (vgl. 3.11). Die melodisch-prosodischen Eigenschaften inklusive der nachfolgenden Rhythmik-Analysen beschränken sich im Rahmen dieser Arbeit auf eine rein deskriptive Darstellung. Die Beobachtungen dienen als Referenzgrundlage für Nachfolgeuntersuchungen innerhalb der Kohortenstudie.

4.4 Rhythmusanalyse in doppelböigen Vokanten

Wie in Kapitel 4.3 erwähnt, wurde zusätzlich eine Rhythmusanalyse bei doppelböigen Vokanten vergleichend zwischen beiden Gruppen durchgeführt.

Dazu erfolgten zwei Ansätze innerhalb der Doppelbogenstrukturen:

- (1) die Untersuchung der Längenverhältnisse zwischen Bogen 1 (LB1) und 2 (LB2) und
- (2) die Untersuchung der Bogenform-Kombinationen.

4.4.1 Längen-Verhältnisse LB1 zu LB2 und Rhythmik-Charakteristika

Im Folgenden werden die Ergebnisse der Analyse der Längenverhältnisse des ersten Bogens zum zweiten innerhalb der doppelböigen Vokanten dargestellt. In Analogie zu doppelsilbigen Äußerungen späterer vokaler Entwicklungsphasen werden diese auch den in der Linguistik üblichen Betonungsmustern zugeordnet (vgl. 3.10). Hierbei wurde zwischen trochäischen und jambischen Mustern und dem sogenannten Levelstress unterschieden. Die Zuordnung der Melodien zu den drei aufgeführten Mustern erfolgte basierend auf den Längenverhältnissen der beiden Bögen und ist in Kapitel 3.10 beschrieben.

Die ermittelten absoluten und relativen Häufigkeiten (Signallevel) der Rhythmik-Charakteristika sind in Tabelle 20 dargestellt.

Tabelle 20: Absolute und relative Häufigkeiten der Rhythmikmuster in doppelböigen Melodien beider Gruppen

Gruppe	Rhythmikmuster					
	Trochäus		Levelstress		Jambus	
	abs	rel [%]	abs	rel [%]	abs	rel [%]
HI	19	50,0	14	36,8	5	13,2
KG	28	29,2	38	39,6	28	29,2

Tabelle 20 zeigt ein relativ ausgewogenes Verteilungsbild der Rhythmikmuster innerhalb der Kontrollgruppe. Es ist keine Tendenz in Richtung Jamben oder Trochäen erkennbar.

4.4.2 Bogenform-Kombinationen in doppelböigen Vokanten

Die Variable Alpha (α_1 für Bogen 1, α_2 für Bogen 2) markiert den relativen Zeitpunkt des Auftretens des f_0 -Maximums zur jeweiligen Bogenlänge (vgl. 3.10). Im Folgenden werden hierbei drei Bogenformen unterschieden („fallend“, „symmetrisch“, „steigend“ – siehe Kap. 3.10). Die entsprechenden Bogenformen sind jeweils zum besseren Verständnis in Tabelle 21 grafisch skizziert.

Tabelle 21: Bogenform-Kombinationen bei doppelböigen Vokanten in beiden Gruppen

Gruppe	α_2		Fallend		Symmetrisch		Steigend		Σ	
	α_1								Bogen1	
HI	Fallend		1	2,6%	9	23,7%	10	26,3%	20	52,6%
	Symmetrisch		2	5,3%	10	26,3%	2	5,3%	14	36,8%
	Steigend		2	5,3%	0	0,0%	2	5,3%	4	10,6%
Σ Bogen2 HI			5	13,2%	19	50%	14	36,8%	38	100%
KG	Fallend		17	17,7%	18	18,8%	33	34,4%	68	70,8%
	Symmetrisch		0	0,0%	10	10,4%	11	11,5%	21	21,9%
	Steigend		2	2,1%	0	0,0%	5	5,2%	7	7,3%
Σ Bogen2 KG			19	19,8%	28	29,2%	49	51,0%	96	100%

Tabelle 21 zeigt, dass die fallende Bogenform, bei der die Betonung am Bogenanfang liegt, mit einem Anteil von 52,6% bei der HI-Gruppe und 70,8% bei der KG-Gruppe die dominierende Konturform des ersten Bogens in beiden Gruppen darstellt. Betrachtet man das Konturverhalten des zweiten Bogens in den doppelböigen Vokanten-Strukturen, so zeigt sich das gegenteilige Bild: Wie in Tabelle 21 gezeigt, stellt der Anteil der „fallenden“ Strukturen im zweiten Bogen mit 13,2% bei der HI-Gruppe und 19,8% bei der KG-Gruppe den geringsten gegenüber den anderen beiden Konturformen dar.

5 Diskussion

5.1 Anteil der Vokanten im Lautrepertoire

Nach Buder und Oller werden Vokanten („Quasivowels“) vor allem in den ersten beiden Lebensmonaten produziert, bleiben aber über den gesamten vokalen Entwicklungszeitraum der ersten 12 Monate präsent (Buder et al. 2013). Genauere Angaben über Auftrittshäufigkeiten finden sich in der Literatur nicht. Ab dem zweiten Lebensmonat werden Vokanten mit ersten konsonantähnlichen Elementen kombiniert und erzeugen das typische Klangbild der sogenannten Gurr-Laute. Vokanten bleiben auch im marginalen Babbeln die typischen Vokalelemente oft in Kombination mit Konsonanten. Erst im Verlauf des zweiten Lebenshalbjahres treten volle Vokale im Rahmen von Silben des kanonischen Babbelns auf. Vokanten sind nicht einfach nur eine primitive Form späterer voller Vokale, sondern ein regulärer Bestandteil des Lautrepertoires über das gesamte erste Lebenshalbjahr.

In der vorliegenden Untersuchung wurde ein Anteil der Vokanten am Gesamtlautrepertoire von durchschnittlich 20,8% bei den Hörbehinderten, sowie von 15,7% bei der Gruppe der normalhörenden Säuglinge festgestellt. Wobei dieser Anteil innerhalb der Probanden und Probandinnen sehr unterschiedlich war und Werte zwischen 7% und bis zu 47% bei den Gesunden annahm. Bei den Hörbehinderten variierten die Anteile zwischen 11% und bis zu 70%. Bei zwei Säuglingen der HI-Gruppe wurde ein Anteil von >50% beobachtet (vgl. Tabelle 6, Kap. 3.8). Vergleichsdaten gibt es in der Literatur dazu nicht.

In einer Untersuchung von Clement von jeweils drei hochgradig sensorineural hörgeschädigten Säuglingen und drei gesunden Säuglingen im Alter von 2,5 – 5,5 Monaten wurden Unterschiede für verschiedene Vokalisationstypen im Lautrepertoire erfasst (Clement 2004). Aus methodischen Gründen können daraus keine Angaben bezüglich der Vokantenhäufigkeit entnommen werden. Es zeigte sich aber, dass >79% der Äußerungen im Repertoire der HI-Säuglinge einfache Phonationen

ohne artikulatorische Bewegung waren, während es bei den Gesunden Säuglingen nur 55% waren.

Der Grund dafür, dass der Vokantenanteil zwischen den Säuglingen in der vorliegenden Arbeit stark schwankt, hat möglicherweise auch bei den Kohortenprobanden mit der nur einmaligen Aufzeichnung ihrer Laute zu tun. Es wurde dadurch nur eine kurze Momentaufnahme gemacht, die den tatsächlich vom Säugling erzeugten Vokantenanteil möglicherweise nicht repräsentativ reflektiert. Die Auftrittshäufigkeit bei den Probanden mit Archivdaten ist repräsentativer und spricht gegen den Einwand: Die Auftrittshäufigkeiten der drei längsschnittlich aufgezeichneten Säuglinge betrug AB03: 12%, AB02: 14%, und AB01: 18%. Bei den vier hörbehinderten Säuglingen war sie ähnlich (BB04: 11%, BB01: 12%, BB02: 17% und BB03: 23%). Das zeigt, dass die einfache Häufigkeit, mit der Vokanten erzeugt werden aufgrund der individuellen Variation kein zuverlässiger Indikator für die Frühdiagnostik darstellt. Dennoch zeigten die hochgradig hörgestörten Säuglinge eine Tendenz zu einer höheren Auftrittshäufigkeit was zeigt, dass sich hier bereits Verzögerungen in der vorliegenden Arbeit andeuten können.

Zusammengefasst zeigte sich, dass die Vokantenauftrittshäufigkeit allein kein Indikator für eine vorsprachliche Entwicklungsstörung im ersten Lebenshalbjahr zu sein scheint. Das ähnliche Vorkommen in beiden Gruppen könnte bedeuten, dass Vokanten, die in relativ neutraler Vokaltraktposition erzeugt werden, eine Art „Reset-Variante“ darstellen, die der Säugling regelmäßig erzeugt, um diese Einstellung mit der orofazialen Einstellung zur Produktion voller Vokale und/oder konsonant-ähnlicher Elemente zu kontrastieren. Ob sich die beiden Gruppen in den übrigen Lautkategorien unterscheiden, wird gegenwärtig im Projekt nachuntersucht.

5.2 Formantanalyse

Die Anatomie des Vokaltraktes eines Säuglings unterscheidet sich maßgeblich von der eines Erwachsenen. Die vereinfachte Betrachtung eines infantilen Vokaltraktes als verkleinerte Form eines Ausgewachsenen wird der komplexen Entwicklung nicht gerecht (Kent und Murray 1982; Lieberman et al. 1972; Vorperian et al. 2005). Die Zunge ist flach und füllt fast den gesamten Mund aus, was ihre Beweglichkeit noch stark einschränkt und die Artikulation anfänglich limitiert. Insbesondere in den ersten 18 Lebensmonaten vollstreckt sich eine auserpägte Veränderung der Lage- und Größenbeziehungen im Bereich des Vokaltraktes; der Kehlkopf sinkt ab und der Vokaltrakt verlängert sich (Vorperian et al. 2005). Zusätzlich werden die feinmotorischen und neuronalen Regelmechanismen bei der Lautproduktion reifer. Diese Veränderungen spiegeln sich in der Lautbildung und deren Variationsmöglichkeiten, insbesondere im Bereich der artikulatorischen Möglichkeiten, wider. Während sich in den ersten Lebensmonaten die Lautbildung primär auf rein laryngealer Ebene bewegt, entwickeln sich erst mit zunehmendem Alter und damit zunehmender Reife des Vokaltraktes des Säuglings verstärkt supralaryngeale artikulatorische Modifikationsmöglichkeiten in der Lautproduktion.

Im Bereich der Vokalerzeugung lässt sich diese Entwicklung in den Resonanzeigenschaften des Vokaltraktes, die durch Variation der supralaryngealen Strukturen und Bewegungen beeinflusst werden, beobachten. Man spricht hierbei von sogenannten Formanten, die durch Resonanzbildung im Vokaltrakt entstehen und allerdings auch durch Verfeinerung der sprechmotorischen Kontrolle und der Etablierung interner Repräsentationen der Vokaltraktkonfiguration beeinflusst werden. (Vorperian und Kent 2007). Durch Messung der Formantwerte lassen sich Rückschlüsse auf die Größe und Form des Vokaltraktes ziehen und Aussagen über die Stellung der Weichgewebe während der Lautbildung und der Länge des Vokaltraktes treffen.

Bei Vokanten handelt es sich, wie im Kapitel 3.7.2 definiert, um vokalähnliche egressive Vokalisationen mit und ohne laryngeale Konstriktionen, aber ohne zusätzliche supralaryngeale Artikulationen. Vokalähnlich deshalb, weil die für die Vokale typischen dominanteren Obertonbereiche, also resonanzverstärkte Teiltöne (Formanten) noch nicht vom Säugling intentional und systematisch vokalspezifisch erzeugt werden.

Vokanten ähneln daher den Vokalen nur bedingt in den Erzeugungsmechanismen. Bei den Vokanten ist der Vokaltrakt definitionsgemäß in „neutraler“ Position, d.h. Resonanzfrequenzen werden nicht intentional variiert. Dadurch sollten sich in Vokanten primär die durch die Vokaltraktgeometrie bedingten Resonanzfrequenzen finden. Bei intentionaler Erzeugung „voller“ Vokale werden spezifische Resonanzfrequenzen orofazial „eingestellt“. Im Verlauf der Entwicklung über die ersten Lebensmonate erlaubt die anatomischen Veränderung des Vokaltraktes, insbesondere das Absinken des Larynx und die gleichzeitige neurophysiologische Reifung beteiligter Kontrollsysteme intentionale Veränderungen der Vokaltrakteinstellungen. Diese erzeugen unterschiedliche sonore Klänge ohne bei Säuglingen schon die, für die deutschen Vokale typischen Resonanzfrequenzen (der ersten beiden Formanten F1 und F2) aufzuweisen.

Mit zunehmender Vokaltraktlänge nehmen die Resonanzfrequenzwerte ab, womit eine Resonanzfrequenzverschiebung mit dem Alter und geschlechtsspezifisch zu beobachten ist (Masapollo et al. 2016; Vorperian und Kent 2007). Zur Beschreibung der Verhältnisse wurde das sogenannte Vokaldreieck eingeführt.

Das Vokaldreieck bildet sich im Formant-Diagramm aus den ersten beiden Resonanzfrequenzen der Vokale [a], [i], [u], die mit ihren Formantwerten die Grenzbereiche der Resonanzmöglichkeiten und damit den Vokalraum darstellen. Jeder Punkt im Diagramm codiert somit für eine bestimmte Vokalqualität, die allein durch ihre Resonanzfrequenzen einem Vokal zugeordnet werden kann (Masapollo et al. 2016).

Vokanten nutzen in ihren Resonanzmöglichkeiten noch nicht das volle Potenzial des Vokalraumes aus. Der eingeschränkte Resonanzraum gibt dieser Lautkategorie auch ihren Namen. In der Regel zeigen sie breitbandige Resonanzen geringer Amplitude, primär unter 1200Hz, ab dem 4. Monat auch über 1200Hz (Jun Murai 1964; Nakazima 1962). Bedingt durch ihre anatomischen und neuromotorischen Einschränkungen bezüglich supralaryngealer Mechanismen, zeigen Neugeborene und junge Säuglinge eine geringere Variationsmöglichkeit der Formanten. Mit zunehmender Reifung zeigt sich eine leichte Veränderung der Formantwerte hin zu kleineren Frequenzwerten (Kuhl und Meltzoff 1996) und eine signifikante Abnahme ihrer Bandbreite ab dem 4. Monat

(Robb et al. 1997). Die Abnahme der Bandbreite wird einerseits auf die Veränderung der Gewebezusammensetzung (Viskosität) im Vokaltrakt zurückgeführt, andererseits auf die Abnahme der Nasalität. Letzteres ist mit der Reifung der velo-pharyngealen Funktion verbunden (Sasaki et al. 1977).

Einfache, isoliert auftretende Vokanten zeigen relativ stationäre „Formanten“, die die Resonanzeigenschaften der Vokaltrakt-Geometrie widerspiegeln. Die „neutrale“ Vokaltrakt-Position, die für diese Vokanten charakteristisch ist, weicht jedoch nach Buder et al. (2013) etwas vom sogenannten „e-Schwa-Vokal“ ab.

Die für Formanten relativ geringen Standardabweichungen der hier berechneten F1- und F2-Werte belegen die Unbeweglichkeit der Resonanzen und damit die Neutralposition des Vokaltrakts bei der Erzeugung der Vokanten. Dies zeigt die korrekte Identifizierung der Vokanten im Lautrepertoire vor der Bestimmung ihrer melodischen Merkmale, die das Ziel der Arbeit war.

Vokanten haben typischerweise häufig noch einen hohen Nasalitätscharakter. Dadurch entsteht eine Dominanz der Energie in den niedrigen Frequenzen und man findet häufig keine Harmonischen oberhalb von 2,5 kHz, was auf einen leicht offenen velopharyngealen Übergang hindeutet. Die Nasalität dämpft die Gesamt-Lautenergie und verbreitert dadurch die Formantbandweiten. Diese einfachen Vokanten werden häufig sehr leise produziert, womit die Energie oftmals nicht viel über dem Hintergrundgeräuschlevel liegt, was zur gedämpften Energieverteilung führt und die tiefen Frequenzen betont. Dies erschwert manchmal ihre spektographische Identifikation (Buder et al. 2013), sodass auch in den hier berichteten Daten einige Werte fehlen. Es wurden nur Werte im Vokaldreieck eingezeichnet bei denen nachvollziehbare F1 und F2 Berechnungen vorlagen. Dies wurde interaktiv kontrolliert (vgl. Kap. 3.8) und es gab demzufolge eine geringfügige Differenz zwischen den in 4.1, Abbildung 8 im Vokalraumdiagramm verwendeten Daten und allen weiteren Analysen. Um den vierten Monat herum wird die Nasalität geringer (Oller 1986). Während die weniger nasalisierten Laute Frequenzen oberhalb 2000Hz (F2) annehmen, haben Thom et al. und Hsu et al. beobachtet, dass Säuglingslaute mit nasalem Einschlag die Formantwerte etwas senken (Hsu et al. 2000; Thom et al. 2006). Der Rückgang der Nasalität wird als

Indikator für das Erreichen eines velopharyngealen Verschlusses angesehen (ibid.). Allerdings haben Thom et al. gezeigt, dass ein vollständiger velopharyngealer Verschluss bis zum sechsten Lebensmonat nicht zu beobachten war. Nasalisierte Laute treten also regelmäßig im Untersuchungszeitraum auf (Thom et al. 2006). Dies bestätigt die Befunde der Arbeit.

Im vorliegenden Lautmaterial wurde ebenfalls eine Absenkung des F2-Formanten durch die Nasalität beobachtet. Oller beschreibt, dass die Vokanten häufig mit einem fast geschlossenen Mund erzeugt werden. Das könnte auch verursacht haben, dass sich unser Hauptcluster der F1- und F2-Werte tendenziell eher an den Formantwerten des späteren Vokals /u/ befanden (Oller 2000).

Rendall et al. haben den Formantraum von Affenlauten (Paviane), sogenannte Grunzlaute, untersucht (Rendall et al. 2005). Obwohl diese Laute in der Grundfrequenz deutlich unter der, der hier untersuchten Säuglinge liegen, zeichnen sich diese Laute durch harmonische Muster aus, die dem der Vokanten entsprechen. Die gefundenen Formantwerte lagen tendenziell ebenfalls bei dem späteren /u/ Vokal. Ähnliche Befunde beschreiben (Boë et al. 2019). Zwar handelt es sich bei Pavianen nicht um Menschenaffen, deren Vokaltrakt häufig mit dem von Säuglingen direkt verglichen wird (Lieberman et al. 1972), aber er dürfte dem der Schimpansen nicht sehr unähnlich sein. Die hier gefundene Tendenz der Resonanzen Richtung /u/-Vokal spiegelt die unreifere Situation der supralaryngealen Artikulationsmechanismen wider.

Die marginale Tendenz einer F1-F2-Werteverteilung in Richtung des späteren Vokals /a/ wurde vor allem in der Kontrollgruppe beobachtet. Erklärbar ist diese Beobachtung eventuell dadurch, dass der Luftstrom durch die Lippen bei /u/- ähnlichen Vokanten stärker fühlbar ist als bei /a/-ähnlichen Vokanten, die durch eine stärkere Mundöffnung erzeugt werden.

So wie die Formantwerte in das Vokaldreieck (Abbildung 8) eingetragen wurden, erhält man für die einzelnen Vokantenkategorien eine Verteilung, die der Beschreibung nach Eigenschaften der Mundöffnung reflektiert. F1 nimmt mit steigendem Öffnungswinkel des Mundes, also mit tiefer liegender Zunge zu. Bei geschlossener Mundstellung, also mit höherer Zungenlage ist der erste Formant erniedrigt. Das entspricht dem Trend der dargestellten Vokanten-Formanten, insbesondere in der HI-Gruppe.

Das Zentrum des akustischen Vokalraums der Säuglinge wird von Kent und Murray mit Werten von 0,5-1,5kHz und 1,8-3,8kHz für F1 und F2 angegeben (Kent und Murray 1982). Die Werte wurden mittels visueller Auswertung von Breitband-Spektralanalysen ermittelt. Diese Formantfrequenzen in Kombination mit der erhöhten Grundfrequenz bei Säuglingslauten, die Werte bis 2000Hz annehmen kann, führt dazu, dass Säuglingslaute vom erwachsenen Ohr als hochfrequent wahrgenommen werden (Kent et al. 1987). Bauer und Kent geben an, dass Formantfrequenzen für die Vokanten und Vokale von Säuglingen ungefähr doppelt so hohe Werte haben, wie die für erwachsene männliche Sprecher (Bauer und Kent 1987). Der Frequenzbereich für jeden Formanten erhöht sich kontinuierlich mit der Entwicklung im ersten Lebensjahr. Im Alter von 6 Monaten sind die mit F1 bei 0,5-1,7kHz und F2 bei 1,6-3,8kHz noch mit denen des 3. Monats vergleichbar (Kent und Murray 1982).

Die von uns im vergleichbaren Alter ermittelten Werte stimmen sehr gut mit den in der Literatur berichteten Formantwerten überein (Buhr 1980; Kuhl und Meltzoff 1996; Polka et al. 2014; Serkhane et al. 2008; Vorperian und Kent 2007).

Wermke et al. (2002) haben gezeigt, dass bereits ab der 9. postnatalen Lebenswoche typische Abstimmungsphänomene zwischen der Melodie und den ersten beiden Resonanzfrequenzen beobachtet werden können. Auch Kempf hat dies in ihrer Dissertation durch Untersuchungen an Säuglingen mit LKGS-Spalten gefunden (Kempf 2010). Vom menschlichen Säugling werden die Resonanzeigenschaften des Vokaltraktes ab dem zweiten Monat vorsichtig ausgetestet, anfänglich wahrscheinlich häufig auch eher zufällig „angesprochen“. Inwieweit diese Abstimmungsmechanismen durch Hörleistungen beeinflusst werden, indem sie bei Einschränkungen der Hörfähigkeit zum Beispiel später oder lokal langsamer erfolgen könnten, ist bisher nicht untersucht. Vorstellbar wäre es aber, dass die Abstimmung zwischen laryngealen Mechanismen (Melodie) und Resonanzmechanismen im Vokaltrakt (Formanten) nicht allein durch propriozeptive Rückkopplungsmechanismen realisierbar ist, sondern einer funktionierenden auditiven Rückkopplung bedarf (Keck 2018).

Die hier untersuchten Vokanten sind in einem Alterszeitraum aufgezeichnet worden der nach Oller als Expansionsstadium (Expansion stage) (Oller und Kent 2001), nach Zlatin M (1975) als exploratorisches phonatorisches Verhalten (Exploratory phonatic

behaviour) und nach Stark (1980) als vokalisches Spielen (vocal play) bezeichnet wird (Stark 1980).

Die Variabilität (Standardabweichung) und der Frequenzumfang (Min / Max) waren bei der Kontrollgruppe größer. Das bedeutet, dass das Gehörte beim Ausprobieren der Klangeigenschaften tatsächlich auch in der eigenen spielerischen Lautproduktion, selbst bei diesen einfachen hier untersuchten Vokanten reflektiert wird (Boero 2015). Säuglinge zeigen im dritten Lebensmonat bereits eine reife spektrale Auflösungsfähigkeit für Frequenzen unter 4000Hz (Cabrera und Werner 2017). Das bedeutet, dass sie die Unterschiede in Klangeigenschaften bei Resonanzvariationen auch tatsächlich wahrnehmen können.

Das spielerische Ausprobieren in der stimmlichen Klangerzeugung ist aber bei Vokanten nur auf laryngeale Produktionsprozesse beschränkt. Das erklärt auch, warum sich die F1-F2-Werte der Vokanten beider Gruppen im Vokaldreieck überdecken.

Die Vokanten selbst beschreiben also eine Lautkategorie, die noch keine typischen Vokalformanten aufweisen, sondern Resonanzfrequenzen, die in erster Linie von den geometrischen Konstellationen des Vokaltraktes und deren zufällige Variation, z.B. durch laryngeale Begleitartikulation oder, Variation der Mundöffnung, abhängen. Wie das LAM-Modell von Esling postuliert, gibt es anfänglich kaum egressive Vokalisationen, die nicht durch laryngeale Konstriktionen begleitet sind (Esling et al. 2019). Ob dies tatsächlich der Fall ist, bleibt noch durch nachfolgende Studien zu beweisen. Dennoch konnte auch bei den Vokanten ohne laryngeale Konstriktionen im Lautrepertoire der hier untersuchten Säuglinge häufig eine sehr kurze (< 100ms) laryngeale Konstriktion am Anfang oder Ende der Vokanten beobachtet werden. Inwieweit diese Phänomene zur Beeinflussung der gemessenen Resonanzfrequenzen (Formanten) geführt haben, ähnlich co-artikulativen Einflüssen in späteren Konsonant-Vokal-Verbindungen des Sprechens, und damit die Variation der F1/F2-Werte im Vokalraum mit erklärt, kann hier nicht entschieden werden.

5.3 Melodiestrukturanalyse

Es liegt bisher nur eine Vorläuferarbeit vor, die Melodiemodulationen in Vokanten untersucht und beschrieben hat (Blum). Diese Arbeit wurde ebenfalls im Rahmen der Kohortenstudie angefertigt (Kap. 3.1). Blum hat gezeigt, dass Vokanten bezüglich ihrer Melodiestruktur (Gesamtgestalt) zwar vorwiegend aus einbögigen Formen bestehen, komplexe Formen aber ebenfalls vorkommen können. Daher wurde hier zunächst untersucht welche Melodiestrukturkategorien die zu analysierenden Vokanten aufwiesen.

Die Analyse der Melodiestruktur ist Bestandteil der Routinediagnostik am Zentrum für vorsprachliche Entwicklung und Entwicklungsstörungen (ZVES). Es hat sich gezeigt, dass die Komplexität der Melodiestruktur im zweiten/dritten Lebensmonat ein Frühindikator für eine Sprachentwicklungsverzögerung sein könnte (Wermke et al. 2011; Wermke, Robb, Schluter 2021; Wermke, Sereschk et al. 2021). Lange Zeit wurde die Melodiestruktur nur im spontanen Säuglingsweinen ermittelt, da entsprechende Vergleichsstudien für die Nichtschrei-Vokalisationen nicht vorlagen. Wermke, Robb, Schluter (2021) haben jetzt ein erstes statistisches Modell zur erwartenden Auftrittshäufigkeit komplexer Melodien in Komfortlautvokalisationen normalhörender Säuglinge veröffentlicht. Das statistische Modell prädiziert einen Anteil komplexer Melodiestrukturen von ca. 30% im Alter von 60 Tagen in Komfortvokalisationen. Dieser Anteil steigt über die folgenden Wochen und erreicht im fünften Lebensmonat einen Wert von ca. 50%. Danach fällt der Wert bis zum 180. Tag auf ungefähr 35% leicht ab.

Mit diesem Modell wurde belegt, dass das von Wermke und Mende postulierte Melodie-Entwicklungs-Modell (MD-Modell) nicht nur für Schrei-Vokalisationen sondern auch für Vokalisationen in anderen Affektzuständen und Kommunikationssituationen gilt. Damit ist anzunehmen, dass die Melodiekomplexität in Komfortvokalisationen ebenfalls ein aussagekräftiger Risikomarker für spätere Spracherwerbsstörungen, vermittelt über Störungen im Prosodieerwerb, sein könnte. Dieser Beleg steht allerdings noch aus, da Längsschnittstudien bisher nicht vorliegen. Die Vermutung, dass die Melodiekomplexität ein möglicher Risikoindikator für eine verzögerte sprachliche Entwicklung sein könnte, wird durch die Befunde vorliegender Arbeit unterstützt. Die Argumente dafür werden nachfolgend diskutiert.

Die Kontrollgruppe zeigte einen nahezu doppelt so hohen Komplexitätsgrad der Vokantenmelodien im Untersuchungszeitraum im Vergleich zu dem der HI-Gruppe (vgl. Kap 4.2). Dies lässt vermuten, dass das angeborene Melodieentwicklungsprogramm sowohl einer kinästhetischen als auch einer auditiven Rückkopplung bedarf, um sich optimal zu entfalten. Die Vibrationsempfindungen hochgradig sensorineural hörgeschädigter Säuglinge können nur eine eingeschränkte Rückkopplung liefern, den fehlenden auditiven Input aber nicht vollständig ersetzen. Die ungestörte Entwicklung der melodisch-rhythmischen Eigenschaften basiert auf neuronalen Regelkreisen, die sensorische Informationen aus beiden Sinnesmodalitäten bedürfen. Die Beobachtung von doppel- und vereinzelt auftretenden noch komplexeren Bogenstrukturen bei der Gruppe HI und segmentierten Formen kann durch die Bedeutung der haptischen Sensorik bei der Regulation erklärt werden.

In der wegweisenden Arbeit von Kent und Murray vor 40 Jahren wurden bereits Melodiestructuren analysiert, wenn auch in einfacher Weise durch auditive Transkription (Kent und Murray 1982). Bei der Untersuchung vokalischer Säuglingslaute wurde zwischen einfachen und doppelböigen sowie komplexen Melodiestructuren mit 2 oder mehr Grundformen, unterschieden. Dabei zeigte sich im Alter von 3 und 6 Monaten ein deutliches Überwiegen der einfach strukturierten Vokalisationen mit einböigen steigend-fallenden, oder nur fallenden bzw. steigenden Konturen und relativ frequenzstabilen, flachen Melodien. Komplexere Strukturen wurden von dem Autor nur in Einzelfällen beobachtet, wie es auch hier der Fall war. Doppelböige Melodien traten deutlich häufiger auf als komplexere Formen. Es wurde eine relative Zunahme doppelböiger Melodien vom dritten bis sechsten Monat beobachtet. Damit konnten die Ergebnisse von Blum reproduziert werden.

Der in der KG-Gruppe beobachtete Anteil an doppelböigen (15%) und komplexeren (10%) Strukturen liegt allerdings deutlich unter dem von Wermke, Robb, Schluter (2021) in diesem Alter postulierten für Komfortlaute. Das ist damit zu erklären, dass hier nur Vokanten berücksichtigt sind, während in der genannten Studie alle Vokalisationsarten eingeschlossen waren. Das bedeutet, dass Vokanten seltener komplexe Melodiekonturen aufweisen als andere Vokalisationen des Lautrepertoires.

Das könnte einerseits mit der zeitlichen Dauer zusammenhängen – Vokanten sind vergleichsweise kurz. Der hier beobachtete Anteil an sehr kurzen Vokanten betrug ca. 20% mit einem leicht höheren Anteil in der HI-Gruppe (vgl. Tabelle 11, Kap. 4.2). Diese Kurzlaute, so hat sich auch hier wieder bestätigt, sind ein wichtiger Teil des Lautrepertoires und nicht etwa nur Produkte rein vegetativer Ereignisse. Buder et al. zählen diese Lautkategorie bereits ab dem ersten Lebensmonat zu den sogenannten „Protophonen“, den aus ihrer Sicht sprechrelevanten Lautkategorien (Buder et al. 2013). Die während der audiovisuellen Analysen im Zusammenhang mit den Begleitprotokollen gemachten Erfahrungen bestätigen dies. Daher wurden auch die Kurzlaute bei der Analyse der temporalen Eigenschaften partiell mitberücksichtigt.

Andererseits könnte das seltenere Vorkommen komplexer Melodiestructuren in Vokanten aber auch darauf hindeuten, dass offenbar Rückkopplungsmechanismen des Vokaltraktes auf den Larynx aktiv an der Erzeugung komplexerer Melodien beteiligt sind. Bei Vokanten sind die Rückkopplungsmechanismen durch die neutrale Vokaltraktposition weitaus geringer und einfacher als bei anderen Vokalisationstypen.

Die Melodieentwicklung in vorsprachlichen Lauten verläuft nach dem MD-Modell (Melody-Development-Model) von Wermke und Mende von einfachen, einbögigen Strukturen zu komplexen Strukturen (Wermke und Mende 2009). Blum hat dies für Vokanten bereits gezeigt, fand aber auch eine Entwicklungsverzögerung bei der HI-Gruppe. Nach dem MD-Modell stellen doppelböigige und komplexere Strukturen, die aus Kombinationen von Einzelbausteinen in Form einfacher Bögen gebildet werden, eine höhere Entwicklungsstufe dar. Da Vokanten sehr einfache phonatorische Expressionen sind, könnte der Anteil komplexer melodischer Strukturen hier deutlich niedriger sein. Für die weiterführende Analyse im Gesamtprojekt wird daher ein erhöhter Komplexitätsgrad für das gesamte Lautrepertoire der normalhörenden Säuglinge postuliert. Diese Auswertungen laufen gegenwärtig (Arnold in Vorbereitung).

In komplexen melodischen Strukturen können die Einzelbögen entweder direkt aufeinanderfolgen oder nach einem Bogen durch vollständige oder unvollständige Pausen auch unterbrochen sein (Segmentierungspausen). Solche innermelodischen Pausen entstehen durch Konstriktionen im laryngealen Artikulator (Esling 2005), wie in

Kapitel 1 beschrieben. Solche Unterbrechungen können an beliebiger Stelle in der Melodie erfolgen und auch in ihrer Anzahl pro Vokalisation variieren. Nach dem LAM-Modell handelt es sich bei diesen rhythmischen Mechanismen um Vorstufen der späteren Silbenbildung (Esling et al. 2019). Dies ist im Einleitungsteil (vgl. Kap. 1) bereits ausgeführt worden. Im Rahmen der vorliegenden Arbeit wurde auch die Frage gestellt, ob die Erzeugung laryngealer Konstriktionen in Vokanten durch den Verlust der auditiven Rückkopplung beeinflusst werden könnten. Es stellte sich also die Frage, ob sich die segmentierten Melodiestructuren in den Vokanten der Patientengruppe in ihrer Häufigkeit von derjenigen der Vokanten der Kontrollgruppe unterscheiden.

Es wurde ein Anteil segmentierter Melodiestructuren in Vokanten von rund 9,3% in der Kontrollgruppe identifiziert. In den Vokanten der HI-Gruppe waren es rund 6,5%. Das könnte darauf hindeuten, dass das Fehlen der auditiven Rückkopplung bei den hochgradig sensori-neural geschädigten Säuglingen zu einer selteneren Erzeugung laryngealer Konstriktionen in Vokanten führt. Allerdings war die Stichprobe sehr klein und die Vermutung wird durch die Arbeit von Schenk (2022) entkräftet. Zwar wurden in ihrer Arbeit hohe interindividuelle Unterschiede in der Häufigkeit laryngealer Konstriktionsphänomene festgestellt, ein signifikanter Unterschied zwischen normalhörenden und hörgeschädigten Säuglingen wurde allerdings nicht gezeigt. Trotz fehlendem auditiven Feedback unter den sensori-neural gestörten untersuchten Säuglingen, liefert Schenk mit dem Phänomen der propriozeptiven Rückkopplung einen möglichen Erklärungsansatz für ihre Beobachtungen (Schenk 2022).

Die Rhythmisierung durch laryngeale Konstriktionen und Schwingungspausen der Stimmlippen, die der zeitlichen Strukturierung des Phonationsstroms dienen, wird also vermutlich nicht primär durch das Hören beeinflusst, sondern vorrangig durch propriozeptive Rückkopplung.

Die signifikant höhere Komplexität der Melodiestructur in Vokanten lieferte ein Unterscheidungsmerkmal zwischen den Gruppen. In der weiteren Datenauswertung im Rahmen der Kohortenstudie muss analysiert werden, ob es eine signifikante Korrelation zwischen der Hörleistung und der Melodiekomplexität gibt. Danach ließe sich dann entscheiden, ob die Melodiekomplexität ein Kandidat für die Frühdiagnose einer Hörstörung im klinischen Setting sein könnte.

5.4 Dauer von Vokanten

Zlatin et al. haben eine allgemeine Zunahme der Länge von Komfortvokalisationen über die ersten 24 Wochen gemessen und haben beobachtet, dass Säuglinge in diesem Alter verschiedene Vokalisationslängen systematisch variieren (Zlatin et al. 1977).

Für die Vokanten selbst sind keine analogen Angaben in der Literatur zu finden; sie werden allgemein nur als „kurz“ beschrieben (Bauer und Kent 1987; Kent und Murray 1982). Bauer und Kent geben eine Dauer von weniger als 400ms im Ergebnis der Auswertung früherer Arbeiten für die überwiegende Mehrheit der Vokalisationen im ersten Lebensjahr an (Buhr 1980; Delack et al. 1978). Die Messungen in der vorliegenden Arbeit bestätigen diese Angabe: Es fand sich eine rechtsschiefe Verteilung, mit relativ mehr Vokanten unterhalb von 400ms (vgl. Abbildung 12 und Abbildung 13). Der Variationsbereich der Vokantenlängen lag in der vorliegenden Arbeit zwischen 154ms und 1330ms (HI-Gruppe) bzw. 1631ms (KG-Gruppe). Die längeren Vokanten der Kontrollgruppe sind vermutlich in erster Linie darauf zurückzuführen, dass komplexere Melodiestrukturen häufiger vorkamen. Der identische Minimalwert in beiden Gruppen ist gegebenenfalls auf die Messmethodik zurückzuführen, da Laute unter einer Länge von 150ms nicht mit einbezogen wurden. Das hatte den Grund intentional erzeugte Kurzlaute von rein vegetativen Geräuschen physiologischer Ereignisse zu separieren (vgl. Kap. 3.9.1).

Auch die Dauer der am häufigsten aufgetretenden Melodiestrukturen, der einbögigen Vokanten, unterschied sich nicht zwischen den Gruppen. Hierbei zeigten sich vergleichbare Medianwerte mit nur 30ms Unterschied. Im Ergebnis der Arbeit stellt der Autor zunächst die Hypothese auf, dass es keine Unterschiede in den Vokantenlängen zwischen normalhörenden und sensori-neural hörgestörten Säuglingen im Alter von 3-6 Monaten gibt. Das bedeutet, dass die zeitliche Organisation bei der Erzeugung von Vokanten eine neurophysiologisch robuste Eigenschaft ist, die kaum durch auditive Rückkopplungseffekte beeinflusst wird. Lediglich indirekt, über die Komplexität der Melodiestruktur sind diese Effekte wirksam, wie bereits erwähnt.

Es bleibt festzuhalten, dass es sich bei Vokanten um einfache Strukturvarianten handelt, die, definitionsgemäß, kaum intentionalen Regelungen unterliegen. Damit sind hier wahrscheinlich atem- und stimmphysiologische Mechanismen primäre Einflussgrößen.

Die Befunde unserer Studie haben dennoch gezeigt, dass sowohl bei der Beurteilung sämtlicher untersuchter Vokanten, unter Berücksichtigung aller Strukturformen, mit $p < 0,001$ für das Alter bzw. $p = 0,037$ für die Gruppe, als auch für die einbögigen Vokanten, mit jeweils $p < 0,001$ bezüglich beider Parameter, ein signifikanter Alters- und Gruppeneffekt bestand. Obwohl kortikale Bereiche an der Lautproduktion in diesem Alter beteiligt sind, wird für die Vokanten eher eine Regelung auf subkortikaler Ebene vermutet, wie bei den Ruflauten nicht-humaner Primaten (Ploog 2004).

Bei vergleichbarer mittlerer Vokantenlänge zeigte sich im Rahmen unserer Untersuchungen in beiden Gruppen allerdings eine relativ hohe Variabilität, wie sie sowohl von Bauer und Kent, als auch von Oller beobachtet wurden (Bauer und Kent 1987; Oller 1980). Diese Variation betrifft alle Laute in der Explorationsphase (4.-6. Monat), aber auch die Vokanten, wie hier erstmalig belegt wurde. Die Form der melodischen Gestalt realisiert sich zeitlich variabel.

In der Arbeit wurde eine strenge Selektion nach Vokanten vorgenommen und es ist erstaunlich, dass ein vergleichbarer Variationsbereich in den Melodielängen beobachtet wurde, wie er von anderen Autoren für das gesamte Komfortvokalisations-Repertoire beschrieben ist. Hsu et al. (2000) haben die Vokalisationslängen von 13 (5w) gesunden Säuglingen in den ersten 6 Lebensmonaten untersucht und dabei herausgefunden, dass silbenähnliche Laute länger waren als vokalische Laute (Hsu et al. 2000). Die mittlere Dauer silbenartiger Vokalisationen einfacher Struktur betrug im dritten Lebensmonat: 410ms; im vierten: 480ms und im fünften, wie im sechsten Lebensmonat: 340ms. In einer vergleichbaren Studie von Gregory et al. (2018) wurden vier australische Säuglinge, ebenfalls in den ersten 6 Lebensmonaten untersucht. Nach Auswertung von insgesamt 6309 Vokalisationen fand man im fünften Lebensmonat im

Mittel Vokalisationsdauern pro Kind von: 360ms; 416ms; 457ms bzw. 434ms (Gregory et al. 2018).

Dies könnte auf universale temporale, im Gehirn verankerte, Mechanismen der Zeitorganisation von Lautproduktionen zurückgeführt werden. Auch, ohne ausgeprägte auditive Rückkopplungsmöglichkeit im Fall einer hochgradig sensorineuralen Hörstörung wird die Lautlänge explorativ variiert. Allerdings ist auch nicht ausgeschlossen, dass die auch bei hochgradig sensorineural geschädigten Säuglingen vorhandene Fähigkeit zur Wahrnehmung von Vibrationen bzw. im Einzelfall auch eine gewisse Resthörigkeit zu solchen Längenvariationen führen können. Dies kann anhand der vorliegenden Daten nicht geklärt werden.

Interpretiert werden diese Längenunterschiede von Clement mit dem Alter von dem Beginn der Explorationsphase um den vierten Monat herum (Clement 2004). Die Autorin betont, dass in ihrer Untersuchung alle Säuglinge, auch die der HI-Gruppe, in der Lage waren lange Vokalisationen (> 2000ms) zu erzeugen. In der Regel werden die längeren Vokalisationen nach dem dritten Lebensmonat durch anatomische und physiologische Faktoren erklärt, zum Beispiel die Veränderung der Rippenknochen (Langlois et al. 1980). Durch das Wachstum der Rippen und entsprechender muskulärer und neurophysiologischer Entwicklung, können Säuglinge die Dauer erzeugter Vokalisationen besser kontrollieren, indem sie den subglottalen Luftdruck ab diesem Alter besser regulieren können. Diese Vermutung in der Literatur ist aus Sicht des Autors zu relativieren, da bereits eine erhebliche Längenkontrolle im spontanen Weinen Neugeborener von Sereschk belegt wurde (Sereschk 2019; Wermke, Sereschk et al. 2021). Es ist allerdings fraglich, ob diese signifikanten Unterschiede der Vokantenlänge tatsächlich inhaltlich von Bedeutung sind, um normalhörende Säuglinge von hochgradig sensori-neural hörgestörten Probanden und Probandinnen zu differenzieren.

Es gibt nur eine wissenschaftliche Studie, die sich mit Untersuchungen temporaler Eigenschaften von Vokanten befasst und die zu einem gegenteiligen Befund kommt

(Binos et al. 2021). Binos et al. (2021) gehen davon aus, dass die Vokantenlängen normalhörender Probanden und Probandinnen länger sind als die der höreingeschränkten. Allerdings wurden von den Autoren lediglich zwei Kinder analysiert, die zudem ein höheres Alter aufwiesen, als die Säuglinge hier. Die Arbeit lässt sich also nur bedingt mit den Ergebnissen unserer Studie vergleichen.

Bloom hat die Vokalisationslänge von 80 drei Monate alten Säuglingen untersucht und dabei herausgefunden, dass silbenähnliche Laute länger waren als vokalische Laute, unabhängig vom sozialen Kontext, in dem sie erzeugt wurden (Bloom 1989). Die mittlere Dauer silbenartiger Vokalisationen betrug 1,22s (Spannweite 0,68s – 2,43s), die der vokalischen Laute lag bei 0,87s (Spannweite: 0,43-2,0).

Schnelle auditorische Verarbeitungsmechanismen und die schnelle Erkennung von zeitlichen Veränderungen sind für die Sprech- und Sprachentwicklung von hoher Bedeutung, insbesondere im ersten Lebensjahr. Es gibt Hinweise darauf, dass die Unterscheidungsfähigkeit für schnelle Frequenzveränderungen insbesondere in der linken Hemisphäre bei Säuglingen im Alter von 4 Monaten präsentiert ist (Musacchia et al. 2013). Grundlage dieser Testung waren Tonunterscheidungsaufgaben von 400Hz absolut (800-1200Hz) mit einer Intervallpause von 30 - 70ms im Wechsel mit dem Standardreiz (800Hz). Es zeigte sich generell eine größere Reaktion der rechten Hemisphäre, die linke Hemisphäre reagiert nur auf die sehr schnellen Frequenzwechsel. Die linke Hemisphäre scheint also insbesondere auf schnelle Tonwechsel spezialisiert zu sein, wie sie auch für das spätere Sprechen erforderlich sind. Die rechte Seite dagegen ist mehr für die Tonhöhen- und Melodiewahrnehmung verantwortlich. Auswirkungen von Lateralitätsunterschieden der Hemisphäre sind bei der Vokantenerzeugung wahrscheinlich eher marginal.

5.5 Melodisch-Rhythmische Eigenschaften in doppelböigen Vokanten

Die Melodiekontur ist das wichtigste Wahrnehmungselement sprachlicher und musikalischer Stimuli für Säuglinge (Armbrüster 2020; Iversen et al. 2008; Wermke und Mende 2009; Zatorre und Baum 2012). In der vorliegenden Arbeit ist die Melodiekontur in Form der Melodiestrukturkategorien untersucht worden. Dabei zeigte sich, dass 15% aller analysierten Vokanten der Kontrollgruppe bereits die erste Komplexifikationsstufe in der Melodieentwicklung, die doppelböige Melodiekontur aufwiesen. In der HI-Gruppe waren dies immerhin auch 7% (Kap. 0). Eine Analyse dieser 2B-Strukturen erlaubte es auch, einfache rhythmische Eigenschaften zu untersuchen (Kap. 4.4). Dazu wurde die Dauer des ersten bzw. zweiten Bogens ermittelt und das Verhältnis beider gebildet (Kap. 3.10).

Bei der Untersuchung der rhythmischen Variationsfähigkeit fiel auf, dass beide Gruppen einen etwa gleich hohen Anteil (37-40%) an isochronen Mustern hatten (Levelstress). Ein wesentlicher Unterschied bestand in der Häufigkeit trochäischer Muster, also Rhythmen des Musters („lang-kurz“) und dem gegenteiligen Muster „kurz-lang“ (jambisch). Dieses Muster trat in den Vokanten mit doppelböiger Struktur (2B) deutlich häufiger in der HI-Gruppe auf. Nach der Theorie der akustischen Gestaltwahrnehmung ist das häufigere Muster das kurz-lang Muster (Lerdahl und Jackendoff 1983). Bereits vor mehr als einhundert Jahren haben Forscher zwei universale Prinzipien postuliert, die bei der Wahrnehmung einfacher Tonsequenzen (Melodie), also beim akustischen Gruppieren eine Rolle spielen. Das erste Prinzip besagt, dass die Intensität den Beginn einer neuen Gruppe in der Wahrnehmung markiert, also die Lautheit ein Segmentierungssignal liefert. Das wurde in dieser Arbeit nicht untersucht. Relevant ist hier das zweite universale Prinzip, dass eine verlängerte Tonsequenz tendenziell das Ende einer akustischen Gruppe markiert (Bolton 1984; Woodrow 1909). Dieses Prinzip ist bereits bei Säuglingen zu beobachten: Trainor and Adams (2000) haben Erwachsene und 8 Monate alte Säuglinge untersucht, indem sie Pausen (silent gaps) zwischen aufeinanderfolgenden Tönen gesetzt haben und damit zeigen konnten, dass

sowohl Säuglinge als auch Erwachsene die längeren Töne als Markierung für das Ende der wahrgenommenen Töne markierten.

Das häufigere Auftreten des Musters „lang-kurz“ in doppelböigen Vokanten der HI-Gruppe könnte eine physiologische Ursache haben: Es wird ein Bogen erzeugt und danach, wenn die Luft langsam knapp wird, nur noch ein kurzer drangehängt. Bei der KG-Gruppe kommt das trochäische „lang-kurz“ Muster zwar auch vor, aber in derselben Häufigkeit wie das jambische „kurz-lang“ Muster (29%). In der HI-Gruppe gab es nur 13% des jambischen Musters in doppelböigen Vokanten. Dieses Ergebnis wird vorläufig so interpretiert, dass die angeborene „innere“ akustische Gruppierungspräferenz einer auditiven Rückkopplung bedarf.

Obwohl das Wahrnehmen von akustischen Ereignissen nicht nur auf temporalen Eigenschaften basiert, sondern auch Intensität- und Tonhöhenvariation einschließt, scheint die temporale Segmentierungsfähigkeit eine der ersten Eigenschaften zu sein, die Säuglinge erlernen (Trainor und Adams 2000; Weber et al. 2004).

Die Frage danach, ob es sich bei der Unterscheidung von (kurz-lang / lang-kurz) rhythmischen Mustern tatsächlich um eine innere universale Wahrnehmungseigenschaft des Gehirns handelt oder ob sie von der akustischen Umgebung abhängig ist, wurde von Iversen et al. neu gestellt (Iversen et al. 2008). Die Autoren fanden, dass diese temporale Gruppierungseigenschaft stark von der sprachlichen Umgebung abhängt. Muttersprachliche englische und japanische erwachsene Sprecher wurden bezüglich der Gruppierungswahrnehmung einfacher rhythmischer Tonsequenzen untersucht. Es zeigten sich unterschiedliche Muster bei der Wahrnehmung, die die Autoren dahingehend deuteten, dass die Unterscheidung kurz-lang nicht universal sei, sondern durch Erfahrung geformt wird.

Die, in der vorliegenden Arbeit untersuchten normalhörenden Säuglinge haben eine ein- bis viermonatige Erfahrung durch die Umgebungssprache. Bei einem Großteil der hier untersuchten Säuglinge handelt es sich nicht um monolingual deutschsprechende Eltern, die Kinder haben also unterschiedlichste rhythmische Erfahrungen gemacht.

Damit lässt sich auch erklären, dass zu etwa gleichen Teilen trochäische (lang-kurze) Rhythmen bzw. jambische (kurz-lang) in der Kontrollgruppe beobachtet wurden. Neben dem Grundmuster der Isochronie werden also rhythmische Variationen selbst in den einfachen Vokanten produziert. Das Vorherrschen des trochäischen (50%) und vergleichsweise geringen Vorkommen des jambischen Musters (13%) bei den hochgradig sensorineural Hörgeschädigten, könnte mit der fehlenden auditorischen Erfahrung mit der Umgebungssprache zusammenhängen. Das von diesen Säuglingen bevorzugte trochäische Muster kann aufgrund der fehlenden Hörfähigkeit nicht auf den Einfluss der Umgebungssprache zurückgeführt werden. Vielmehr scheint hier eine atemphysiologische Erklärung sinnvoll, indem es sozusagen keiner intentionalen rhythmischen Variation bedarf, eine lange phonatorische Expression zu äußern und am Ende, wenn die Luft knapp ist, noch einen kürzeren Laut in der Entspannungsphase nachzufügen - also das Muster lang-kurz zu äußern. Der Autor nimmt hier keine intentionale Rhythmusvariation an und sieht in den rhythmischen temporalen Eigenschaften der 2B-Strukturen ein möglicherweise geeignetes Merkmal, um in weiteren Analysen im Rahmen der gesamten Kohortenstudie robuste Unterscheidungen zwischen Vokalisationen normalhörender und hörgestörter Säuglinge treffen zu können.

Die hier untersuchten Rhythmen gehören zu den einfachsten, die man vorsprachlich untersuchen kann. Komplexere, wie die Variation der Längen vom Beginn einer Vokalisation zur nachfolgenden in der Originalaufnahme, sogenannte Inter-Onset-Intervalle (Rachor 2022), in den aufgezeichneten Lautsequenzen der Säuglinge liefern signifikante Unterschiede in Abhängigkeit von der Hörleistung und bestätigen die Bedeutung temporaler Größen in der vorsprachlichen Diagnostik. Das basiert darauf, dass die funktionierende auditive Rückkopplung beim Säugling eine Grundvoraussetzung dafür ist, in den eigenen Lauten rhythmische Variationen zu erzeugen und sich somit an den gehörten Umgebungssprachrhythmus anzupassen (Entrainment).

Zur Erklärung der jambischen Rhythmusbeobachtungen der Kontrollgruppe scheint das Phänomen der Verlängerung der Endsilbe („Final Syllable Lengthening“) möglicherweise eine Rolle zu spielen. Untersuchungen der universalen infant-directed-speech (IDS), haben gezeigt, dass das Verlängerungsphänomen der finalen Silbe ein typisches Merkmal im IDS ist. Kondaurova et al. (2012) haben Vokale im IDS von Müttern mit hochgradig sensorineural hörgeschädigten und normalhörenden Babys im Alter von 11 Monaten untersucht (Kondaurova et al. 2012). Die Mütter beider Säuglingsgruppen hyperartikulierten den akustischen Vokalraum und verlängerten die Vokaldauer bei direkter Anrede des Kindes (IDS), unabhängig von dessen Hörfähigkeiten. Die Dauer der Laute an sich unterschied sich ebenfalls nicht zwischen der Gruppe (Morgan et al. 1987).

Das Phänomen der Verlängerung der letzten Silbe findet man auch in tonalen Sprachen (Tang et al. 2017). Tang et al. vermuten, dass die Mütter im IDS die Dauer der Wörter und Silben in der Endstellung besonders dehnen und damit betonen, um die Wortsegmentierungen für die Säuglinge zu erleichtern. Das steht in Übereinstimmung mit der Beobachtung, dass Kinder, wenn sie Wörter erlernen, häufig Objekte nur mit der finalen Silbe benennen. Beispiel anstelle von Banane sagen sie „ne“ oder „nane“. Der Anteil der jambischen Muster der normalhörenden Säuglinge erklärt sich dann damit, dass diese bereits umfangreich durch das IDS ihrer Eltern beeinflusst worden sind.

5.6 Formkombinationen in doppelböigen Vokanten

Es wurde die Form (Kontur) des ersten bzw. zweiten Melodiebogens bestimmt und danach Formkombinationen ermittelt (Kap. 4.4.2). Das Aufzählen der Kombinationsmuster lieferte in 71% aller doppelböigen Vokanten der Kontrollgruppe ein fallendes Muster im ersten Bogen. Bei der HI-Gruppe begannen nur 53% der Vokanten mit einer fallenden Kontur.

Mampe et al. untersuchten und verglichen bereits 2009 das Intonationsmuster von einböigen Schreien deutscher und französischer gesunder Säuglinge und stellten dabei unterschiedliche Betonungsmuster anhand ihrer Melodieformen fest. Es stellte sich heraus, dass deutsche Säuglinge bereits unmittelbar nach Geburt eine deutliche Präferenz zu anfangsbetonten Schreilauten aufweisen, während es bei französischen Säuglingen eher endbetonte sind (Mampe et al. 2009). Auch im Rahmen der vorliegenden Auswertungen zeigte sich die Dominanz dieser fallenden Muster mit 71% insbesondere unter den gesunden Säuglingen, was die Befunde der Studie von Mampe et al. bestätigt. Die Präferenz zum muttersprachlichen Intonationsmuster setzt jedoch eine Hörprägung voraus, die bereits intrauterin stattfindet, wie in der Arbeit von Mampe et al. (2019) gezeigt werden konnte. Bei den sensorineural hörgestörten Säuglingen unserer Kohortengruppe, ließe sich eine solche Prägung nur durch somatosensorische Empfindungen erklären. Diese könnte sich dann postnatal weiter manifestieren. Die Tatsache, dass mehr fallende Formen in der Kontrollgruppe beobachtet wurden, spricht gegen die Annahme, dass es sich hier um einen rein atemphysiologischen Effekt handelt. So argumentieren auch Mampe et al. für die steigenden Muster der französischen Neugeborenen; die nur möglich sind, wenn die Melodieform gegen den physiologischen Intensitätsabfall geregelt werden kann. Eine bessere Erklärung ist anhand dieser 38 Laute jedoch nicht möglich.

Untersuchungen bezüglich Betonungseigenschaften doppelböiger Vokanten liegen in der Literatur bislang nicht vor. Kottmann (2022) beschäftigte sich jedoch in ihrer Dissertationsarbeit mit dem Betonungsschwerpunkt innerhalb ein- bzw. zweiböiger Lautstrukturen unterschiedlicher Vokalisationstypen anhand ihres Melodieverlaufs

bzw. ihrer Grundfrequenz. Ihre Untersuchungen umfassten 1289 doppelböjige Laute im Längsschnitt von 10 gesunden normalhörenden Säuglingen über die ersten sechs Monate. Kottmann stellte im Rahmen Ihrer Forschungsergebnisse mit rund 22% ein Vorherrschen symmetrischer Melodieverläufe in jeweils beiden Bögen fest (Kottmann 2022). Diese Beobachtungen decken sich grob mit den festgestellten 26% bei den sensorineural hörgestörten Säuglingen dieser Kohorte. In der KG-Gruppe waren es lediglich rund 10%. Es gilt dabei jedoch zu beachten, dass es sich bei Kottmanns Untersuchungen sowohl um Babel- als auch Schrei- und Übergangslaute handelte. Die Ergebnisse sind somit nicht direkt vergleichbar und ein Erklärungsansatz hierfür wäre aufgrund der geringen vorliegenden Datenlage zum jetzigen Zeitpunkt eher spekulativ.

Interessant sind jedoch die kontrastierenden Muster fallend-steigend und steigend-fallend. Die fallend-steigende Kombination fand sich bei 34% der doppelböjigen Vokanten der Kontrollgruppe und 26% der HI-Gruppe. Kottmann gibt dem Muster aus fallenden und steigenden Melodieformen in ihrer Arbeit einen Namen und bezeichnet es als „Flügelform“ aufgrund ihrer visuellen Darstellung im Melodiediagramm. In ihrer Arbeit war dieses mit 11% über die ersten 6 Monate in allen Vokalisationstypen zusammen vertreten. Das umgekehrte Muster (steigend-fallend) der beiden Bogenformen in der 2B-Kombination fand sich im Rahmen vorliegender Untersuchungen in nur 2% (KG) bzw. 5% (HI) der Fälle. Kottmann bestätigt diese Beobachtung und ermittelt ebenso einen Anteil von nur rund 2% dieser selteneren Melodiekontur.

Der Autor vermutet auch hier atemphysiologische Hintergründe für seine Beobachtungen, erhebt jedoch keinen tiefergehenden Interpretationsanspruch aufgrund des geringen Datensatzes. Die unabhängige Feststellung dieser Melodieeigenschaften in beiden Forschungsarbeiten ist dennoch bemerkenswert. Um den Betonungseigenschaften zweiböjiger Lautstrukturen, potenziell auch in Hinblick auf Intensitäts- und Lautstärke-Eigenschaften (Davis et al. 2000), nachzugehen, sind weitere Auswertungen in der Kohortenstudie erforderlich. Entsprechende Analysen sind bereits in Arbeit.

6 Limitationen

In der vorliegenden Arbeit wurden die Vokalisationseigenschaften von Vokanten erstmalig objektiv bestimmt. Das Problem bzw. die Herausforderung der zusätzlichen subjektiven Wahrnehmung und Interpretation der einzelnen Lautbildungen blieb dennoch bestehen. Vor diesem Hintergrund wurden die Daten stets zu zweit und mehrfach untersucht und in unsicheren Fällen gegenseitig abgeglichen, um die Möglichkeit einer individuellen Fehleinschätzung zu minimieren.

Bei der Vermessung der Vokantenlängen könnten insofern Unsicherheiten bestehen, als dass nicht am Zeitsignal ermessen wurde, sondern an der Grundfrequenzkontur. Der Autor schätzt, dass dadurch unter Umständen Unsicherheiten in der Größenordnung von 10-80ms entstanden sein können, die bei der Interpretation der statistischen Ergebnisse berücksichtigt wurden. Eine weitere Limitation der Untersuchung besteht darin, dass Querschnitts- und Längsschnitt-Daten gemischt wurden. Die Längsschnitt-Daten nehmen einen sehr großen Anteil der Daten ein, stammen aber nur von 7 der insgesamt 26 Säuglinge. Auch vor dem Hintergrund, dass dies durch die Anwendung gemischter Modelle in der statistischen Analyse berücksichtigt wurde, kann ein Einfluss dieses Designs nicht ausgeschlossen werden. Leider war dies aufgrund der zur Verfügung stehenden Daten zum Zeitpunkt der hier durchgeführten Arbeit nicht vermeidbar. Weiterhin konnte der Einfluss der Umgebungssprache durch teilweise unsichere Dokumentationen nicht zuverlässig analysiert werden. Zum Teil fehlten die Angaben in den Dokumentationen zu den Kohortenteilnehmern.

Die im Längsschnitt untersuchten Säuglinge hatten im Untersuchungszeitraum zeitweilig Hörgeräte aktiviert. Das könnte Einfluss auf die Melodieeigenschaften der Vokanten haben, der allerdings als sehr gering eingeschätzt wird: Aufgrund dessen, dass es sich hier um hochgradig sensorineural Hörgestörte handelte, wird davon ausgegangen, dass die Versorgung mit Hörgeräten keine relevante Verbesserung der Hörleistung mit sich brachte.

7 Zusammenfassung

In der vorliegenden Arbeit wurden spontane Lautäußerungen (Komfortlaute) von normalhörenden Säuglingen und Säuglinge mit einer hochgradig sensorineuralen Hörstörung untersucht. Dabei handelte es sich um Teilnehmer und Teilnehmerinnen einer Kohortenstudie, die gemeinsam vom Comprehensive Hearing Center (CHC) und dem Zentrum für vorsprachliche Entwicklung und Entwicklungsstörung (ZVES) durchgeführt wird. Diese wurden mit Probanden und Probandinnen vorhergehender Längsschnitt-Studien ergänzt. Ziel der Arbeit war es die einfachsten Vokalisationstypen aus dem Komfortlautrepertoire zu analysieren, um der Frage nachzugehen, ob es bereits bei rein phonatorisch erzeugten Vokalisationen Unterschiede zwischen den Säuglingen beider Gruppen geben könnte. Dieser Ansatz hatte den Vorteil der Analyse eines sehr homogenen Lautrepertoires vor dem Hintergrund einer hohen intra- und interindividuellen Variabilität in der frühen Sprachentwicklung.

Es wurden rein vokalische Laute, sogenannte Vokanten, objektiv mithilfe signalanalytischer Methoden untersucht. Erstmals sind damit auch Referenzwerte für die Melodielänge von Vokanten über den Untersuchungszeitraum berichtet. Die Eigenschaften dieses typischen Vokalisationstyps junger Säuglinge sind bisher in der Literatur nicht beschrieben und wurden hier erstmalig untersucht.

In der Stichprobe von 8 sensorineural hörgestörten und 18 normalhörenden Säuglingen und insgesamt 1236 Vokanten fanden sich statistisch signifikante Unterschiede (Gemischte Modelle) in der Melodiekontur (Zeitfunktion der Grundfrequenz). Die normalhörenden Säuglinge zeigten einen signifikant höheren Anteil an komplexen Melodiekonturen im Vergleich zu den hochgradig hörgeschädigten Säuglingen. Keine inhaltlich bedeutsamen Unterschiede fanden sich in der Vokalisationslänge der Vokanten. Allerdings zeigte eine einfache rhythmische Analyse doppelböiger Vokanten in beiden Gruppen, dass die fehlende auditorische Erfahrung in der hörgestörten Gruppe zu einer von dem Rhythmus der normalhörenden Säuglinge geäußerten Vokanten abwich. Dies wurde mit der Notwendigkeit der Erfahrung für die akustische Gestaltwahrnehmung begründet. Dabei handelt es sich zunächst nur um eine

vom Autor aufgestellte Hypothese, die durch nachfolgende Studien geprüft werden muss. Insgesamt hat die vorliegende Untersuchung gezeigt, dass es bereits in einem sehr frühen Alter und bei sehr einfachen Lautäußerungen zu Unterschieden in Vokalisationseigenschaften kommt.

8 Literaturverzeichnis

- Armbrüster, L. (2020). *Melodieintervalle als Kenngröße der laryngealen Regelleistung bei der Lautproduktion von Säuglingen ohne orofaziale Spaltbildung in den ersten vier Lebensmonaten*. Dissertation, Universität Würzburg.
- Bahgat, Y. S., Khalil, Y. M., El Maghraby, R. A., El Sayed Mohamed, E. & Elsayed, M. M. A.-N. (2012). The use of overnight pulse oximetry and phoniatrics parameters in the screening protocol of obstructive sleep apnea. *Egyptian Journal of Chest Diseases and Tuberculosis* 61 (4), 459–468. doi:10.1016/j.ejcdt.2012.08.017
- Bauer, H. R. & Kent, R. D. (1987). Acoustic analyses of infant fricative and trill vocalizations. *The Journal of the Acoustical Society of America* 81 (2), 505–511. doi:10.1121/1.394916
- Bauer, H. R. (1988). The ethologic model of phonetic development: I. Phonetic contrast estimators. *Clinical linguistics & phonetics* 2 (4), 347–380. doi:10.3109/02699208808985265
- Biesalski, P. (Hrsg.). (1994). *Phoniatrie - Pädaudiologie. In 2 Bänden* (2., neubearb. u. erw. Aufl.). Stuttgart: Thieme.
- Binos, P., Thodi, C., Vogazianos, P., Psillas, G. & Constantinidis, J. (2021). An acoustic and auditory analysis of vocants in infants with cochlear implants. *Logopedics, phoniatrics, vocology* 46 (1), 28–34. doi:10.1080/14015439.2020.1724325
- Bloom, K. (1989). Duration of early vocal sounds. *Infant Behavior and Development* 12 (2), 245–250. doi:10.1016/0163-6383(89)90011-8
- Blum, D. M. *Vergleichende Analyse melodischer Eigenschaften von frühen Komfortlauten bei gesunden und hochgradig hörgeschädigten Säuglingen*. Dissertation, Universität Würzburg.
- Boë, L.-J., Sawallis, T. R., Fagot, J., Badin, P., Barbier, G., Captier, G., Ménard, L., Heim, J.-L. & Schwartz, J.-L. (2019). Which way to the dawn of speech?: Reanalyzing half a century of debates and data in light of speech science. *Science advances* 5 (12), eaaw3916. doi:10.1126/sciadv.aaw3916
- Boero, D. L. (2015). Sounds of protolanguages: Some preliminary insights from developmental psychology. *Theoria et Historia Scientiarum* 11, 99. doi:10.12775/ths-2014-006
- Bolton, W. F. (1984). "The Language of 1984". *English World-Wide* 5 (2), 297.

- Buder, E. H., Warlaumont, A. S., Oller, D. K., Peter, B. & MacLeod, A. (2013). An acoustic phonetic catalog of prespeech vocalizations from a developmental perspective. *Comprehensive perspectives on child speech development and disorders: Pathways from linguistic theory to clinical practice* 4, 103–134.
- Buhr, R. D. (1980). The emergence of vowels in an infant. *Journal of speech and hearing research* 23 (1), 73–94. doi:10.1044/jshr.2301.73
- Cabrera, L. & Werner, L. (2017). Infants' and Adults' Use of Temporal Cues in Consonant Discrimination. *Ear and Hearing* 38 (4), 497–506. doi:10.1097/AUD.0000000000000422
- Cebulla, M., Hofmann, S. & Shehata-Dieler, W. (2014). Sensitivity of ABR based newborn screening with the MB11 BERAphone®. *International journal of pediatric otorhinolaryngology* 78 (5), 756–761. doi:10.1016/j.ijporl.2014.02.003
- Cebulla, M. & Shehata-Dieler, W. (2012). ABR-based newborn hearing screening with MB11 BERAphone® using an optimized chirp for acoustical stimulation. *International journal of pediatric otorhinolaryngology* 76 (4), 536–543. doi:10.1016/j.ijporl.2012.01.012
- Chesters, J., Baghai-Ravary, L. & Möttönen, R. (2015). The effects of delayed auditory and visual feedback on speech production. *The Journal of the Acoustical Society of America* 137 (2), 873–883. doi:10.1121/1.4906266
- Clement, C. J. (2004). *Development of vocalizations in deaf and normally hearing infants* (LOT, Bd. 100). Utrecht: LOT (Zugl.: Amsterdam, Univ., Diss., 2004).
- Da Silva, D., Lopez, P. & Mantovani, J. C. (2015). Auditory brainstem response in term and preterm infants with neonatal complications: the importance of the sequential evaluation. *International archives of otorhinolaryngology* 19 (2), 161–165. doi:10.1055/s-0034-1378137
- Davis, B. L., MacNeilage, P. F., Matyear, C. L. & Powell, J. K. (2000). Prosodic correlates of stress in babbling: an acoustical study. *Child development* 71 (5), 1258–1270. doi:10.1111/1467-8624.00227
- Delack, J. B., Fowlow, D. J., Waterson, N. & Snow, C. (1978). The ontogenesis of differential vocalizations: Development of prosodic contrastivity during the first year of life. *The Development of Communication (London)*, 93–110.
- Delb, W. & Praetorius, M. (2004). Physiologie und Pathophysiologie der Hörwahrnehmung. *Sprache · Stimme · Gehör* 28 (2), 51–59. doi:10.1055/s-2004-818995

- Eilers, R. E. & Oller, D. K. (1994). Infant vocalizations and the early diagnosis of severe hearing impairment. *The Journal of pediatrics* 124 (2), 199–203. doi:10.1016/s0022-3476(94)70303-5
- Eisenmann, M. S. *Entwicklung, Testung und Anwendung von PRAAT Skripten zur semi-automatischen Vokalisationsanalyse im Rahmen der Vorsprachlichen Diagnostik am ZVES*. Dissertation, Universität Würzburg.
- Esling, J. H. (2005). There Are No Back Vowels: The Laryngeal Articulator Model. *The Canadian Journal of Linguistics / La revue canadienne de linguistique* 50 (1), 13–44. doi:10.1353/cjl.2007.0007
- Esling, J. H., Moisik, S. R., Benner, A. & Crevier-Buchman, L. (Hrsg.). (2019). *Voice quality. The laryngeal articulator model* (Cambridge studies in linguistics, Bd. 162). Cambridge: Cambridge University Press.
- Fant G. (1960). *Acoustic theory of speech production*.
- Fruhmann, M. Ein Modell zur Wahrnehmung der Ausgeprägtheit der Tonhöhe. <https://mediatum.ub.tum.de/doc/618351/618351.pdf>. Zugegriffen: 2. Mai 2022.
- Gregory, A., Tabain, M. & Robb, M. (2018). Duration and Voice Quality of Early Infant Vocalizations. *Journal of speech, language, and hearing research : JSLHR* 61 (7), 1591–1602. doi:10.1044/2018_JSLHR-S-17-0316
- Hsu, H.-C., Fogel, A. & Cooper, R. B. (2000). Infant vocal development during the first 6 months: speech quality and melodic complexity. *Infant and Child Development* 9 (1), 1–16.
- Hsu, H.-C., Fogel, A. & Messinger, D. S. (2001). Infant non-distress vocalization during mother-infant face-to-face interaction. *Infant Behavior and Development* 24 (1), 107–128. doi:10.1016/S0163-6383(01)00061-3
- Iversen, J. R., Patel, A. D. & Ohgushi, K. (2008). Perception of rhythmic grouping depends on auditory experience. *The Journal of the Acoustical Society of America* 124 (4), 2263–2271. doi:10.1121/1.2973189
- Iyer, S. N. & Oller, D. K. (2008). Prelinguistic Vocal Development in Infants with Typical Hearing and Infants with Severe-to-Profound Hearing Loss. *The Volta review* 108 (2), 115–138.
- Jun Murai. (1964). The Sounds of Infants : Their Phonemization and Symbolization. *Studia phonologica* (1964), 3: 17-34.

- Keck, J. (2018). *Pilotuntersuchung zur Anwendbarkeit einer Phonations-Artikulations-Interaktionsanalysemethode zur Charakterisierung artikulatorischer Mechanismen in kanonischen Babbellauten von Säuglingen mit hochgradiger sensorineuraler Hörminderung*. Doctoralthesis, Universität Würzburg.
- Kempf, A. (2010). *Untersuchung des Beginns artikulatorischer Vorgänge in den Lautproduktionen von Säuglingen mit orofazialen Spalten*. Dissertation, Universität Würzburg.
- Kent, R. D. & Murray, A. D. (1982). Acoustic features of infant vocalic utterances at 3, 6, and 9 months. *The Journal of the Acoustical Society of America* 72 (2), 353–365.
- Kent, R. D., Osberger, M. J., Netsell, R. & Hustedde, C. G. (1987). Phonetic development in identical twins differing in auditory function. *The Journal of speech and hearing disorders* 52 (1), 64–75.
- Kent, R. D. & Bauer, H. R. (1985). Vocalizations of one-year-olds. *Journal of child language* 12 (3), 491–526. doi:10.1017/S0305000900006620
- Kondaurova, M. V., Bergeson, T. R. & Dilley, L. C. (2012). Effects of deafness on acoustic characteristics of American English tense/lax vowels in maternal speech to infants. *The Journal of the Acoustical Society of America* 132 (2), 1039–1049. doi:10.1121/1.4728169
- Kottmann, T. (2022). *Längsschnittanalyse von Stimmparametern bei gesunden Säuglingen in den ersten sechs Lebensmonaten [Im Einreichungsprozess]*. Dissertation, Universität Würzburg.
- Kuhl, P. K. & Meltzoff, A. N. (1996). Infant vocalizations in response to speech: vocal imitation and developmental change. *The Journal of the Acoustical Society of America* 100 (4 Pt 1), 2425–2438. doi:10.1121/1.417951
- Langlois, A., Baken R. J. & Wilder, D. N. (1980). Pre-speech respiratory behaviour during the first year of life. *Infant communication: cry and early speech., Texas, Houston:College Hill Press, 1980*.
- Lerdahl, F. & Jackendoff, R. (1983). An Overview of Hierarchical Structure in Music. *Music Perception: An Interdisciplinary Journal* 1 (2), 229–252. doi:10.2307/40285257

- Lieberman, P., Crelin, E. S. & Klatt, D. H. (1972). Phonetic Ability and Related Anatomy of the Newborn and Adult Human, Neanderthal Man, and the Chimpanzee. *American Anthropologist* 74 (3), 287–307. <http://www.jstor.org/stable/671515>.
- Lind, K. (2007). *Untersuchung spektraler und melodischer Eigenschaften von Säuglingsschreien. Eine Longitudinalstudie über die ersten 16 Lebenswochen im Rahmen der Deutschen Sprachentwicklungsstudie*, Charité - Universitätsmedizin Berlin.
- Mampe, B., Friederici, A. D., Christophe, A. & Wermke, K. (2009). Newborns' cry melody is shaped by their native language. *Current biology : CB* 19 (23), 1994–1997. doi:10.1016/j.cub.2009.09.064
- Mampe-Keller, B. A. (2013). *Auswirkungen der pränatalen Sprachumgebung auf die Schreimelodie - Ein Vergleich zwischen Neugeborenen mono- und bilingualer Mütter*. Dissertation, Universität Würzburg.
- Martin, J. A. M. (1981). *Voice, Speech, and Language in the Child: Development and Disorder*. Vienna: Springer Vienna.
- Masapollo, M., Polka, L. & Ménard, L. (2016). When infants talk, infants listen: pre-babbling infants prefer listening to speech with infant vocal properties. *Developmental science* 19 (2), 318–328. doi:10.1111/desc.12298
- Mende, W., Herzel, H. & Wermke, K. (1990). Bifurcations and chaos in newborn infant cries. *Physics Letters A* 145 (8-9), 418–424. doi:10.1016/0375-9601(90)90305-8
- Morgan, J. L., Meier, R. P. & Newport, E. L. (1987). Structural packaging in the input to language learning: Contributions of prosodic and morphological marking of phrases to the acquisition of language. *Cognitive Psychology* 19 (4), 498–550. doi:10.1016/0010-0285(87)90017-X
- Musacchia, G., Choudhury, N. A., Ortiz-Mantilla, S., Realpe-Bonilla, T., Roesler, C. P. & Benasich, A. A. (2013). Oscillatory support for rapid frequency change processing in infants. *Neuropsychologia* 51 (13), 2812–2824. doi:10.1016/j.neuropsychologia.2013.09.006.
- Nakamura, H., Takada, S., Shimabuku, R., Matsuo, M., Matsuo, T. & Negishi, H. (1985). Auditory nerve and brainstem responses in newborn infants with hyperbilirubinemia. *Pediatrics* 75 (4), 703–708.

- Nakazima, S. (1962). A Comparative Study of the Speech Developments of Japanese and American English in Childhood (1) : A Comparison of the Developments of Voices at the Prelinguistic Period. *Studia phonologica* (1962), 2: 27-46 (2), 27-46.
- Newcombe, R. G. (1992). Practical Statistics for Medical Research. *Journal of Epidemiology & Community Health* 46 (5), 549. doi:10.1136/jech.46.5.549-a
- Oller, D. K. (1980). *THE EMERGENCE OF THE SOUNDS OF SPEECH IN INFANCY* (Child Phonology): Elsevier.
- Oller, D. K., Eilers, R. E., Neal, A. R. & Schwartz, H. K. (1999). Precursors to speech in infancy: the prediction of speech and language disorders. *Journal of communication disorders* 32 (4), 223–245. doi:10.1016/s0021-9924(99)00013-1
- Oller, D. K. (1986). Metaphonology and Infant Vocalizations. In *Precursors of Early Speech* (S. 21–35). London: Palgrave Macmillan UK.
- Oller, D. K. (2000). *The Emergence of the Speech Capacity*: Psychology Press.
- Oller, D. K. & Kent, R. D. (2001). The Emergence of the Speech Capacity. *The Journal of the Acoustical Society of America* 110 (3), 1237–1238. doi:10.1121/1.1388001
- Papoušek, M. (2001). *Vom ersten Schrei zum ersten Wort. Anfänge der Sprachentwicklung in der vorsprachlichen Kommunikation* (Aus dem Programm Huber, 3. Nachdr. der 1. Aufl. 1994). Bern: Huber.
- Paul Boersma & David Weenink. (2018). Praat, v. 6.0.40 [Computer program]. <https://www.fon.hum.uva.nl/paul/praat.html>. Zugegriffen: 18. Januar 2022.
- Ploog, D. (2004). Different in Non-Human Primates and Homo sapiens. *The speciation of modern Homo sapiens* 106, 121–135.
- Polka, L., Masapollo, M. & Ménard, L. (2014). Who's talking now? Infants' perception of vowels with infant vocal properties. *Psychological science* 25 (7), 1448–1456. doi:10.1177/0956797614533571
- Rachor, E. (2022). *Vergleichende Untersuchung rhythmischer Eigenschaften von Komfortvokalisationen gesunder und hochgradig hörbeeinträchtigter Säuglinge im Alter von 2 bis fünf Monaten*. Dissertation, Universität Würzburg.
- Rendall, D., Kollias, S., Ney, C. & Lloyd, P. (2005). Pitch (F0) and formant profiles of human vowels and vowel-like baboon grunts: the role of vocalizer body size and voice-acoustic

- allometry. *The Journal of the Acoustical Society of America* 117 (2), 944–955.
doi:10.1121/1.1848011
- Robb, M. P., Chen, Y. & Gilbert, H. R. (1997). Developmental aspects of formant frequency and bandwidth in infants and toddlers. *Folia Phoniatrica et Logopaedica* 49 (2), 88–95.
doi:10.1159/000266442
- Robb, M. P. & Pang-Ching, G. K. (1992). Relative timing characteristics of hearing-impaired speakers. *The Journal of the Acoustical Society of America* 91 (5), 2954–2960.
doi:10.1121/1.402931
- Robb, M. P., Bauer, H. R. & Tyler, A. A. (1994). A quantitative analysis of the single-word stage. *First Language* 14 (42-43), 37–48. doi:10.1177/014272379401404203
- Robb, M. P. & Cacace, A. T. (1995). Estimation of formant frequencies in infant cry. *International journal of pediatric otorhinolaryngology* 32 (1), 57–67. doi:10.1016/0165-5876(94)01112-B
- 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 : JSLHR* 63 (1), 49–58. doi:10.1044/2019_JSLHR-S-19-0205
- Sasaki, C. T., Levine, P. A., Laitman, J. T. & Crelin, E. S. (1977). Postnatal descent of the epiglottis in man. A preliminary report. *Archives of otolaryngology (Chicago, Ill. : 1960)* 103 (3), 169–171. doi:10.1001/archotol.1977.00780200095011
- Schenk, J. (2022). *Vergleich laryngealer Konstriktionsphänomene in Komfortvokalisationen normalhörender und hochgradig hörbeeinträchtigter Säuglinge im 3.-5. Lebensmonat*. Dissertation, Universität Würzburg.
- Schwarz, Y., Kaufman, G. N. & Daniel, S. J. (2017). Newborn hearing screening failure and maternal factors during pregnancy. *International journal of pediatric otorhinolaryngology* 103, 65–70. doi:10.1016/j.ijporl.2017.09.027
- Sereschk, N. (2019). *Temporale Eigenschaften ingressiver und egressiver Phonationsleistungen gesunder Neugeborener*. Dissertation, Universität Würzburg.
- Serkhane, J. E., Schwartz, J.-L. & Bessière, P. (2008). Building a Talking Baby Robot: A Contribution to the Study of Speech Acquisition and Evolution. In P. Bessiere, C. Laugier & R. Siegwart (Hrsg.), *Probabilistic reasoning and decision making in sensory-motor*

- systems* (Springer Tracts in Advanced Robotics, Bd. 46, Bd. 46, S. 329–357). Berlin: Springer.
- Shapiro, S. M. & Nakamura, H. (2001). Bilirubin and the auditory system. *Journal of perinatology : official journal of the California Perinatal Association* 21 Suppl 1, S52-5; discussion S59-62. doi:10.1038/sj.jp.7210635
- 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. *Speech, Language and Hearing* 16 (2), 99–106. doi:10.1179/2050571X13Z.00000000011
- Smith, B. L., Brown-Sweeney, S. & STOEL-GAMMON, C. (1988). A quantitative analysis of reduplicated and variegated babbling. *First Language* 9 (6), 175–189. doi:10.1177/014272378900900605
- Staffeldt, S. (2010). *Einführung in die Phonetik, Phonologie und Graphematik des Deutschen. Ein Leitfaden für den akademischen Unterricht* (Stauffenburg Einführungen, Bd. 21). Tübingen: Stauffenburg.
- Stark, R. E. (1980). STAGES OF SPEECH DEVELOPMENT IN THE FIRST YEAR OF LIFE. In *Child Phonology* (S. 73–92). Elsevier.
- Stipdonk, L. W., Weisglas-Kuperus, N., Franken, M.-C. J., Nasserinejad, K., Dudink, J. & Goedegebure, A. (2016). Auditory brainstem maturation in normal-hearing infants born preterm: a meta-analysis. *Developmental medicine and child neurology* 58 (10), 1009–1015. <https://pubmed.ncbi.nlm.nih.gov/27168415/>.
- Stuart, A. (2020). Effect of delivery mode on neonate auditory brainstem responses to air- and bone-conducted stimuli. *International journal of pediatric otorhinolaryngology* 139, 110423. doi:10.1016/j.ijporl.2020.110423
- Tang, P., Xu Rattanasone, N., Yuen, I. & Demuth, K. (2017). Phonetic enhancement of Mandarin vowels and tones: Infant-directed speech and Lombard speech. *The Journal of the Acoustical Society of America* 142 (2), 493. doi:10.1121/1.4995998
- Thom, S. A., Hoit, J. D., Hixon, T. J. & Smith, A. E. (2006). Velopharyngeal function during vocalization in infants. *The Cleft palate-craniofacial journal : official publication of the American Cleft Palate-Craniofacial Association* 43 (5), 539–546. doi:10.1597/05-113
- Titze, I. R. (1994). *Principles of voice production*. Englewood Cliffs, NJ: Prentice Hall.

- Trainor, L. J. & Adams, B. (2000). Infants' and adults' use of duration and intensity cues in the segmentation of tone patterns. *Perception & psychophysics* 62 (2), 333–340. doi:10.3758/bf03205553
- 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 : JSLHR* 50 (6), 1510–1545. doi:10.1044/1092-4388(2007/104)
- 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. doi:10.1121/1.1835958
- Weber, C., Hahne, A., Friedrich, M. & Friederici, A. D. (2004). Discrimination of word stress in early infant perception: electrophysiological evidence. *Brain research. Cognitive brain research* 18 (2), 149–161. doi:10.1016/j.cogbrainres.2003.10.001
- Wermke, K. (2002). *Untersuchung der Melodieentwicklung im Säuglingsschrei von monozygoten Zwillingen in den ersten 5 Lebensmonaten.*
- Wermke, K., Birr, M., Voelter, C., Shehata-Dieler, W., Jurkutat, A., Wermke, P. & Stellzig-Eisenhauer, A. (2011). Cry melody in 2-month-old infants with and without clefts. *The Cleft palate-craniofacial journal : official publication of the American Cleft Palate-Craniofacial Association* 48 (3), 321–330. doi:10.1597/09-055
- Wermke, K., Hauser, C., Komposch, G. & Stellzig, A. (2002). Spectral analysis of prespeech sounds (spontaneous cries) in infants with unilateral cleft lip and palate (UCLP): a pilot study. *The Cleft palate-craniofacial journal : official publication of the American Cleft Palate-Craniofacial Association* 39 (3), 285–294. doi:10.1597/1545-1569_2002_039_0285_saopss_2.0.co_2
- 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. doi:10.1080/02699200701659243
- Wermke, K. & Mende, W. (2009). Musical elements in human infants' cries: In the beginning is the melody. *Musicae Scientiae* 13 (2_suppl), 151–175. doi:10.1177/1029864909013002081

- Wermke, K., Robb, M. P. & Schluter, P. J. (2021). Melody complexity of infants' cry and non-cry vocalisations increases across the first six months. *Scientific reports* 11 (1), 4137. doi:10.1038/s41598-021-83564-8
- Wermke, K., Sereschk, N., May, V., Salinger, V., Sanchez, M. R., Shehata-Dieler, W. & Wirbelauer, J. (2021). The Vocalist in the Crib: the Flexibility of Respiratory Behaviour During Crying in Healthy Neonates. *Journal of voice : official journal of the Voice Foundation* 35 (1), 94–103. doi:10.1016/j.jvoice.2019.07.004
- Woodrow, H. (1909). *A quantitative study of rhythm; the effect of variations in intensity, rate, and duration, by Herbert Woodrow* (Archives of psychology, ed. by R. S. Woodworth. no. 14). New York, The Science Press.
- Yavarzadeh, F. (2020). *Auftreten laryngealer Konstriktionsphänomene in verschiedenen Vokalisationstypen der ersten 7 Lebensmonate bei Säuglingen ohne Lippen-Kiefer-Gaumen-Segelspalten*. Dissertation, Universität Würzburg.
- Zatorre, R. J. & Baum, S. R. (2012). Musical melody and speech intonation: singing a different tune. *PLoS biology* 10 (7), e1001372. doi:10.1371/journal.pbio.1001372
- Zlatin, Laufer, Marsha, Yoshiyuki & Horii. (1977). Fundamental frequency characteristics of infant non-distress vocalization during the first twenty-four weeks. *Journal of child language* 4 (2), 171–184. doi:10.1017/S0305000900001586
- Zlatin M. (1975). *Preliminary descriptive model of infant vocalization during the first 24 weeks: Primitive syllabification and phonetic exploratory behaviour*.

9 Eigene Publikationen

Wermke, K., **Clad, F.**, Blum, D., Cebulla, M., Shehata-Dieler, W. (2023). Melody of vocants – fixed pattern or shaped by hearing? *Folia Phoniatrica et Logopaedica*

10 Appendix

I. Abkürzungsverzeichnis

Klinik

AEHP	Auditiv evozierte Hirnstammpotenziale	BERA	Brainstream evoked response audiometry
CHC	Comprehensive Hearing Center	HI	Hearing impaired / Patientengruppe
KG	Kontrollgruppe	NHS	Neugeborenen-Hörscreening
PG	Probandengruppe		

Labor

CDAP	Cry-Data-Analysis-Programm	e	egressiver Laut
EFA	Energie-Frequenz-Amplitude	I	Inspirationsgeräusch
ZVES	Zentrum für vorsprachliche Entwicklung und Entwicklungsstörung		

Melodiestrukturen

1B	Einerbogen	2B	Doppelbogen
KL	Kurzlaut	MB	Mehrfachbogen
MCI	Melodie-Komplexitäts-Index	MD-Modell	Melodie-Entwicklungsmodell
SEG	Segmentierung		

Akustische Messparameter

f_0	Grundfrequenz	F1 / F2	Formant1 / Formant2
LB	Bogenlänge	L_{ges}	Melodielänge

Statistik

AIC	Akaike-Informationskriterium	Abs	Absolutwerte
KI	Konfidenzintervall	M	Mittelwert
Max	Maximum	Md	Median
Min	Minimum	N	Anzahl
OG	Obergrenze	OR	Odds Ratio
Rel	Relative Anteile	SD	Standardabweichung
UG	Untergrenze		

Sonstiges

d	engl. „day“ – Alter in Tagen	IDS	Infant directed speech
LAM	Laryngeal Articulator Model	LPC	Linear Predictive Coding

II. Abbildungsverzeichnis

Abbildung 1: Rekrutierungswege der Untersuchungsgruppen (eigene Darstellung)	12
Abbildung 2: Prozess der Datenverarbeitung [Erklärungen in den nachfolgenden Kapiteln]	16
Abbildung 3: Beispiel einer annotierten Aufnahmesequenz im PRAAT-Ausgabefenster.....	19
Abbildung 4: Breitbandspektrogramme zweier Vokanten	26
Abbildung 5: Schematische Darstellung eines Boxplots	34
Abbildung 6: Q-Q-Diagramm der ln-logarithmierten Melodielängen aller Vokanten (oben) und der Residuen der Melodielängen (unten)	36
Abbildung 7: Prädiktor-Residuen-Streudiagramm des gemischten Modells (ohne Interaktionsterm) der Melodielängen als abhängige Variable unter Berücksichtigung aller Vokanten	36
Abbildung 8: Verteilung der Formantfrequenzen F1 und F2 der Vokanten im Vokalraumdiagramm.....	39
Abbildung 9: Boxplots der Formantfrequenzen F1 und F2 der Vokanten der Gruppe der hochgradig sensorineural gestörten Säuglinge HI (N=331 Vokanten) und der Gruppe der normalhörenden Säuglinge (KG) (N=352 Vokanten).....	40
Abbildung 10: Melodiestrukturverteilung der Vokanten der Gruppe HI und KG.....	42
Abbildung 11: Boxplots für die Melodielänge der Vokanten der Kontrollgruppe (KG) und der Gruppe der hörgestörten Säuglinge (HI)	46
Abbildung 12: Längenverteilung der Vokanten der Gruppe HI unabhängig von der Melodiestruktur	46
Abbildung 13: Längenverteilung der Vokanten der Gruppe KG unabhängig von der Melodiestruktur	47
Abbildung 14: Boxplot der Melodielängen für einböigige (1B) Vokanten und Kurzlaute (KL)	50
Abbildung 15: Längenverteilungen der Vokanten mit einfacher Melodiestruktur der Gruppe HI	51
Abbildung 16: Längenverteilungen der Vokanten mit einfacher Melodiestruktur der Gruppe KG	51
Abbildung 17: Boxplot für die Melodielängen doppelböiger Vokanten für die Kontrollgruppe (KG) und die Gruppe der hochgradig sensorineural hörgeschädigten Säuglinge (HI).....	55

III. Tabellenverzeichnis

Tabelle 1: HI-Gruppe (bilaterale hochgradige Hörstörung)	13
Tabelle 2: Kontrollgruppe	14
Tabelle 3: Alter (Wochen) der Säuglinge bei Aufnahme, Anzahl aufgenommener Vokalisationen und Zeitpunkt der Hörgeräteversorgung (unterstrichene Tage) der Gruppe „HI“	15
Tabelle 4: Annotationsvariablen für Einzelereignisse.....	17
Tabelle 5: Vokanten im Vergleich zu anderen beobachteten Vokalisationstypen im analysierten Datenmaterial.....	23
Tabelle 6: Egressive Vokalisationen gesamt und Anteil Vokanten der Kontroll- (KG) und hörgestörten Gruppe (HI).....	25
Tabelle 7: Prosodierelevante Messgrößen	32
Tabelle 8: Deskriptive Statistik der gemittelten Grund- (f_0) und Resonanzfrequenzen (F1 und F2) der Vokanten	40
Tabelle 9: Verallgemeinertes gemischtes Modell zur Untersuchung des Gruppeneffektes auf die pro Vokant gemittelten Werte des ersten (F1) und zweiten Formanten (F2)	41
Tabelle 10: Beobachtete mehrböjige Vokanten über den gesamten Untersuchungszeitraum	42
Tabelle 11: Absolute und relative Häufigkeit beobachteter Strukturkategorien.....	43
Tabelle 12: Verallgemeinertes gemischtes Modell zur Untersuchung des Gruppeneffektes auf die binäre Größe für die Melodiekomplexität	44
Tabelle 13: Melodielänge aller Vokanten ohne Berücksichtigung der Melodiestructur.....	47
Tabelle 14: Deskriptive Statistik der individuellen Messwerte für die Melodielängen	48
Tabelle 15: Bootstrap-Analyse der ln-Logarithmierten Melodielängen aller Vokanten.....	49
Tabelle 16: Deskriptive Statistik für alle vermessenen Melodielängen einfach strukturierter Vokanten gesamt und ohne Berücksichtigung der Kurzlaute (o KL)	52
Tabelle 17: Bootstrap-Analyse der ln-Logarithmierten Melodielängen der Einerbögen (1B)	53
Tabelle 18: Individuelle Auftrittshäufigkeit der Vokanten mit doppelböjigen Melodien.....	54
Tabelle 19: Deskriptive Analyse der Melodielängen doppelböjiger Vokanten.....	55
Tabelle 20: Absolute und relative Häufigkeiten der Rhythmikmuster in doppelböjigen Melodien beider Gruppen	56
Tabelle 21: Bogenform-Kombinationen bei doppelböjigen Vokanten in beiden Gruppen	57

IV. Danksagung

Zunächst möchte ich mich ganz herzlich bei allen bedanken, die diese Forschungsarbeit ermöglicht haben.

Danke an alle Projektteilnehmenden, Probanden und Probandinnen.

Danke an Frau Professorin Shehata-Dieler, für die intensive Zusammenarbeit und die Bereitstellung unserer Aufnahmen am CHC. Danke auch für die Möglichkeit der Projektvorstellung im Rahmen des Forschungstreffens der HNO.

Danke an Frau Professorin Stellzig-Eisenhauer und an Frau Professorin Kathleen Wermke für die Ermöglichung dieser Projektarbeit im ZVES.

Ganz besonders danken möchte ich auch meiner Projektpartnerin Daria Blum. Auch in schwierigen und herausfordernden Phasen habe ich Dich als Leidensgenossin immer sehr geschätzt. Ich bin dankbar, dass uns diese gemeinsame Zeit nach dem Examen verbunden hat. Danke auch für die Weinabende und Deine erfrischende Ehrlichkeit.

Danke auch an alle weiteren Projektteilnehmenden – Jenny, Elena, Franzi, Julia und Sarah und an das ganze Team des ZVES.

Ein weiterer Dank gilt Herrn Peter Wermke. Danke für Ihre kompetente Unterstützung im EDV-Bereich. Danke auch für Ihre Geduld und Ihr ausgeglichenes Wesen.

Meinen allergrößten Dank möchte ich an Frau Professorin Kathleen Wermke richten. Ohne Sie wäre die ZVES-Familie einfach nicht denkbar.

Ihr enormes Fachwissen und Ihre berufliche Leidenschaft beeindruckten mich tief und haben mich immer wieder aufs Neue inspiriert. Besonders gerne werde ich mich immer an Ihre ausgesprochene Warmherzigkeit und Ihre bereichernde Offenheit erinnern.

Danke für diese besondere Zeit.

Danke an Dich, Saskia, für Deine Unterstützung bei allen Herausforderungen.

Danke an meine Geschwister- Ingrid, Marius, Lars, Miriam, Matthias und Camilla. Ich liebe euch.

Danke an meine Eltern.