

Aus dem Zentrum für Zahn-, Mund- und Kiefergesundheit  
der Universität Würzburg

Poliklinik für Kieferorthopädie

Direktorin: Professor Dr. med. dent. Angelika Stellzig-Eisenhauer

**Entwicklung, Testung und Anwendung von PRAAT Skripten zur semi-  
automatischen Vokalisationsanalyse im Rahmen der Vorsprachlichen  
Diagnostik am ZVES**

Inaugural - Dissertation

zur Erlangung der Doktorwürde der

Medizinischen Fakultät

der

Julius-Maximilians-Universität Würzburg

vorgelegt von

Michael Simon Eisenmann

aus Würzburg

Würzburg, Juni 2020

Referentin: Prof. Dr. Kathleen Wermke

Korreferent: Prof. Dr.-Ing. Mario Cebulla

Dekan: Prof. Dr. Matthias Frosch

Tag der mündlichen Prüfung: 17.12.2020

Der Promovend ist Zahnarzt

Meinen Eltern in Dankbarkeit gewidmet

# Inhaltsverzeichnis

Abbildungsverzeichnis.....	VII
Tabellenverzeichnis.....	XI
1 Einleitung .....	1
1.1 Vorsprachliche Diagnostik .....	1
1.2 Zielsetzung und Methodologie .....	10
1.2.1 Zielsetzung .....	10
1.2.2 Methodologie .....	14
2 Material und Methoden.....	17
2.1 Projektdatenbasis und Messgrößen .....	17
2.2 Vergleich semi-automatische und manuelle Projektdatenanalyse .	19
2.3 Beschreibung der Voranalyse der Testsignale .....	20
2.3.1 Testsignale aus der Patientengruppe der Säuglinge und Kinder mit LKGS-Spaltbildung .....	21
2.3.1.1 Testbeispiel bei Gaumenspalte (HSH) ohne eingesetzte Oberkieferplatte .....	22
2.3.1.2 Testsignal für eine Lautsequenz bei Gaumenspalte (HSH) mit eingesetzter Oberkieferplatte .....	23
2.3.1.3 Testsignale für Lautsequenzen bei doppelseitiger LKGS im Sprachalter .....	24
Testsignal für einen Dreiwortsatz .....	24
Testsignal für einen Vierwortsatz .....	26
2.3.1.4 Testsignale aus der Patientengruppe der Säuglinge mit Sagittalnahtsynostosen .....	27
Testsignal einer Lautsequenz bei Sagittalnahtsynostose.....	27
Testsignal einer Lautsequenz bei Sagittalnahtsynostose (präoperativ).....	29

2.3.2	Testsignal für Babbellaute eines Säuglings aus der Kontrollgruppe .....	30
2.3.3	Testsignale aus der Patientengruppe der Säuglinge mit sensorineuraler Hörbeeinträchtigung oder -störung.....	31
	Testsignal einer Sequenz von Weinlauten vor Versorgung mit Hörhilfen.....	31
	Testsignal einer Sequenz von Komfortlauten nach Versorgung mit Hörhilfen .....	33
2.3.4	Testsignale aus der Entwicklungsphase des Kanonischen Babbelns.....	34
2.4	Kurzbeschreibung der Analysesoftware PRAAT.....	36
2.4.1	Grafische Oberfläche PRAAT .....	36
2.4.1.1	Editorumgebung.....	38
2.4.2	Textgrids.....	40
2.4.3	Die PRAAT Skriptsprache.....	40
	Wichtige Konstrukte der Skriptsprache .....	41
	String- und andere Variablen .....	41
	Eingabeformen.....	41
	Loops (Schleifen) .....	41
	Conditionals .....	42
3	Ergebnisse .....	43
3.1	Teil 1: Aufbau und Testung des erarbeiteten Skriptes .....	43
3.1.1	Skript zur Analyse annotierter Serien .....	45
3.1.1.1	Eingabeform.....	46
3.1.1.2	Outputordner.....	48
3.1.1.3	Filecheck Procedure .....	49
3.1.1.4	Stringliste der Dateinamen erstellen .....	51

3.1.1.5	Äußerer Loop .....	52
3.1.1.6	Korrespondierendes Textgridfile laden.....	52
3.1.1.7	Codeblock für die automatische Ermittlung temporaler Kenngößen (Zeitentabelle) .....	54
3.1.1.8	Extrahieren der gewünschten Zeitabschnitte .....	55
3.1.1.9	Grundfrequenz (fo)- und Intensitätsanalyse .....	57
3.1.1.10	Framespezifische Zeitwerte in Variablen speichern.....	61
3.1.1.11	Werte in Tabelle schreiben und anschließend Textdateien speichern .....	62
3.1.2	Skript zur halbautomatischen Annotation von Langserien .....	68
3.1.2.1	Übersicht.....	68
3.1.2.2	Skriptbeschreibung .....	70
3.1.2.3	Eingabeform.....	71
3.1.2.4	Ausgabeordner und Stringliste der Dateinamen erstellen .....	72
3.1.2.5	Sounddateien laden und Textgriddateien erstellen .....	73
3.1.2.6	Überprüfen und/oder ändern der Voreinstellungen durch den Benutzer.....	74
3.1.2.7	Nummerierung der Intervalle.....	77
3.1.2.8	Abspeichern der Textgrids .....	78
3.2	Teil 2: Ergebnisse der Skripttestung .....	80
3.2.1	Ergebnisse der teilautomatischen Annotation.....	80
3.2.1.1	Ergebnisse der skriptbasierten Annotation der Testsignale aus der Patientengruppe der Probanden mit LKGS.....	80
3.2.1.2	Ergebnisse der skriptbasierten Annotation der Testsignale aus der Patientengruppe der Säuglinge mit Sagittalnahtsynostosen .....	88

3.2.1.3	Ergebnisse der skriptbasierten Annotation der Testsignale aus der Patientengruppe der Säuglinge mit sensorineuraler Hörbeeinträchtigung.....	93
3.2.1.4	Ergebnisse der skriptbasierten Annotation der Testsignale aus der Entwicklungsphase des kanonischen Babbeln unauffälliger Säuglinge.....	96
3.2.2	Ermittlung temporaler Eigenschaften der Testsignale mithilfe des Skriptes zur Analyse annotierter Langserien.....	98
3.3	Teil 3: Anwendung des erarbeiteten Skriptes auf Projektdaten ...	101
3.3.1	Vergleich der Melodiestrukturkategorien in Lauten von Säuglingen mit und ohne orofaziale Spaltbildung .....	101
3.3.2	Vergleich der temporalen Eigenschaften in Säuglingslauten der Probandengruppen des Projektes .....	103
3.3.2.1	Temporale Eigenschaften einböiger (1B) Strukturen .....	105
3.3.2.2	Temporale Eigenschaften nichtsegmentierter zweiböiger Strukturen (2B).....	110
3.3.2.3	Temporale Eigenschaften drei- (3B) und mehrböiger (MB) Strukturen.....	114
3.3.2.4	Temporale Eigenschaften einfachsegmentierte (1S) Strukturen .....	117
3.3.2.5	Temporale Eigenschaften zweifachsegmentierter Melodien (2S Strukturen) .....	129
3.3.2.6	Temporale Eigenschaften mehrfachsegmentierter (3S) Strukturen .....	134
3.3.2.7	Temporale Eigenschaften von KS Strukturen .....	137
3.3.2.8	Temporale Eigenschaften der unklaren Strukturen „fuzzy class“ .....	142
3.4	Inspirationslängen.....	145

4	Diskussion.....	147
4.1	Entwicklung und Testung der SAM-Methode.....	147
4.1.1	Vorteile der skriptbasierten Voranalyse .....	148
4.1.2	Bewertung der Ergebnisse der automatischen Annotation .....	150
4.1.2.1	Segmentierungspausen .....	151
4.1.2.2	Einfluss des Signal-Rausch-Abstandes auf das Annotationsergebnis .....	153
4.1.3	Vergleich mit anderen Analyseansätzen und Zukunftsaussichten .....	155
4.1.4	Ermittlung der Frequenzeigenschaften mit dem Skript zur automatischen Analyse annotierter Langserien .....	156
4.2	Temporale Eigenschaften im Methodenvergleich .....	157
4.3	Projektdatenanalyse .....	159
5	Zusammenfassung.....	163
6	Literaturverzeichnis .....	165
	Danksagung .....	
	Lebenslauf.....	
	Eigene wissenschaftliche Tätigkeit.....	
	Anhang.....	i
A 1	Ergänzungen Skriptbeschreibung .....	i
A 1.1	Eingabefenster .....	i
A 1.2	Filecheck Procedure.....	ii
A 1.3	Äußerer Loop .....	iii
A 1.4	Korrespondierendes Textgridfile laden .....	iv
A 1.5	Zeitentabelle .....	v
A 1.6	Werte in Tabelle schreiben .....	vi
A 1.7	Werte in Textdateien abspeichern .....	vi



A 1.8 Überprüfen und/oder ändern der Voreinstellungen durch den Benutzer . .....	vii
A 1.9 Nummerierung der Intervalle .....	viii
A 2 Messwerte der ermittelten Differenzen der Vokalisationslängen (Lges [s]) zwischen SAM und MM-Methode auf Signallevel .....	ix

# Abbildungsverzeichnis

Abbildung 1:	Zeitbereichsdarstellung und Schmalbandspektrum bis 4 kHz (y-Achse) 0-13s (x-Achse) einer Abfolge von Weinlauten und Inspirationsgeräuschen [KAY-CSL (KAY Pentax Medical)] .....	8
Abbildung 2:	Zeitbereichsdarstellung, Schmalbandspektrum bis 4 kHz und Aufteilung einer Langserie von 4,45s [PRAAT (Boersma & Weenink, 2018a)] .....	12
Abbildung 3:	Testsignal A: Zeitbereichsdarstellung und Schmalband-Spektrum .....	22
Abbildung 4:	Testsignal B: Zeitbereichsdarstellung und Schmalband-Spektrum .....	23
Abbildung 5:	Testsignal C: Zeitbereichsdarstellung und Schmalband-Spektrum .....	25
Abbildung 6:	Testsignal D: Zeitbereichsdarstellung und Schmalband-Spektrum .....	26
Abbildung 7:	Testsignal E: Zeitbereichsdarstellung und Schmalbandspektrum .....	28
Abbildung 8:	Testsignal F: Zeitbereichsdarstellung und Schmalbandspektrum .....	29
Abbildung 9:	Testsignal G: Zeitbereichsdarstellung und Schmalbandspektrum .....	30
Abbildung 10:	Testsignal H: Zeitbereichsdarstellung und Schmalband-Spektrum .....	32
Abbildung 11:	Testsignal I: Zeitbereichsdarstellung und Schmalbandspektrum .....	33
Abbildung 12:	Zeitbereichsdarstellung und Schmalbandspektrum einer Sequenz kanonischer Babbelsilben, die durch vollständige Segmentierungen getrennt sind .....	34
Abbildung 13:	Testsignal J: Zeitbereichsdarstellung und Schmalband-Spektrum .....	35
Abbildung 14:	Praat Objects Fenster mit eingeladenen Sound- und Textgriddateien .....	37
Abbildung 15:	links: Praat Picture Window mit Viewport: Fenster mit Skalierung zum Ausgeben von grafischen Darstellungen: hier mit Pitchkonturen; rechts: PRAAT Info Window: einfaches Textfenster zum Ausgeben von Werten oder Texten .....	37
Abbildung 16:	PRAAT Editorumgebung .....	39
Abbildung 17:	Aufbau des erarbeiteten Skriptes .....	44

Abbildung 18:	Grafisches Eingabefenster.....	46
Abbildung 19:	Auswahlfenster für Filecheck procedure: Warnung vor versehentlichem Überschreiben der Dateien .....	49
Abbildung 20:	Extract intervals GUI-Fenster.....	56
Abbildung 21:	Sound to pitch (ac) Eingabefenster.....	60
Abbildung 22:	Übersicht über den Programmumfang des Skriptes zur halbautomatischen Annotation von Langserien .....	69
Abbildung 23:	Auswahlfenster Voreinstellungen (fo-Suchgrenze (pitch) in Hz, Intensitätsschwellenwert (threshold) in dB) .....	74
Abbildung 24:	Testsignal A: automatisch generierter Annotationsvorschlag.....	82
Abbildung 25:	Testsignal A: manuell korrigierte Annotation.....	82
Abbildung 26:	Testsignal B: automatisch generierter Annotationsvorschlag.....	83
Abbildung 27:	Testsignal B: automatisch generierter Annotationsvorschlag nach Anpassung des Schwellenwertes .....	84
Abbildung 28:	Testsignal B nach manueller Korrektur .....	84
Abbildung 29:	Testsignal C: automatisch generierter Annotationsvorschlag.....	86
Abbildung 30:	Testsignal C mit manuell annotierten Wörtern .....	86
Abbildung 31:	Testsignal D: automatisch generierter Annotationsvorschlag.....	87
Abbildung 32:	Testsignal D nach manueller Korrektur der Annotation.....	87
Abbildung 33:	Testsignal E: automatisch generierter Annotationsvorschlag.....	88
Abbildung 34:	Testsignal E: nach manueller Korrektur der Annotation .....	89
Abbildung 35:	Testsignal F: automatisch generierter Annotationsvorschlag.....	90
Abbildung 36:	Testsignal F nach automatischer Nummerierung der Label....	90
Abbildung 37:	Testsignal G: automatisch generierter Annotationsvorschlag.....	92
Abbildung 38:	Testsignal G nach manueller Korrektur der Annotation .....	92
Abbildung 39:	Testsignal H: automatisch generierter Annotationsvorschlag.....	94
Abbildung 40:	Testsignal H nach manueller Korrektur der Annotation.....	94
Abbildung 41:	Testsignal I: automatisch generierter Annotationsvorschlag.....	95

Abbildung 42:	Testsignal I: manuell korrigierte Annotation .....	95
Abbildung 43:	Testsignal J: automatisch generierter Annotationsvorschlag.....	97
Abbildung 44:	Testsignal J: vom Skript durchnummerierte Annotation .....	97
Abbildung 45:	Histogramm der Verteilung der Messdifferenzen für die Testsignale im Methodenvergleich.....	99
Abbildung 46:	Säulendiagramm zum Vergleich der relativen Auftrittshäufigkeiten der Melodiestrukturkategorien in Säuglingslauten der Probanden mit (CLP, blau) und ohne (KG, lila) orofaziale Spaltbildungen .....	102
Abbildung 47:	Boxplotdiagramm der Verteilungseigenschaften der Gesamtlängen (Lges [s]) aller analysierten Vokalisationen (SAM-Methode).....	103
Abbildung 48:	Boxplotdiagramm: Vergleich der Gesamtlängen (Lges [s]) von 1B- Strukturen im Gruppen- und Methodenvergleich .....	105
Abbildung 49:	Boxplotdiagramm der absoluten Zeitmessungsdifferenz zwischen SAM- und MM-Methode für beide Gruppen (Vokalisationstyp 1B) .....	109
Abbildung 50:	Boxplotdiagramm: Vergleich der Gesamtlängen (Lges [s]) von 2B- Strukturen im Gruppen- und Methodenvergleich .....	111
Abbildung 51:	Boxplotdiagramm: Vergleich der Gesamtlängen (Lges [s]) von 3B- und MB- Strukturen im Gruppen- und Methodenvergleich.....	114
Abbildung 52:	Boxplotdiagramm: Vergleich der Gesamtlängen (Lges [s]) von 1S- Strukturen im Gruppen- und Methodenvergleich .....	117
Abbildung 53:	Gruppenunterschiede für die temporalen Messgrößen der Teilzeiten einzelner Strukturelemente im Gruppenvergleich (1S, SAM-Methode) in [s].....	121
Abbildung 54:	Boxplotdiagramm: Gesamtlängen (Lges [s]) aller temporalen Messgrößen für Vokalisationen mit dem melodischen Strukturtyp 1S aus beiden Gruppen.....	125
Abbildung 55:	Boxplotdiagramm: Vergleich der Gesamtlängen (Lges [s]) von 2S- Strukturen im Gruppen- und Methodenvergleich .....	130
Abbildung 56:	Boxplotdiagramm: Vergleich der Gesamtlängen (Lges [s]) von 3S- Strukturen im Gruppen- und Methodenvergleich .....	135
Abbildung 57:	Boxplotdiagramm: Vergleich der Gesamtlängen (Lges [s]) von KS- Strukturen im Gruppen- und Methodenvergleich.....	138
Abbildung 58:	Boxplotdiagramm: Vergleich der Gesamtlängen (Lges [s]) von fuzzy class- Strukturen im Gruppen- und Methodenvergleich.....	143

Abbildung 59: Boxplotdiagramm der Inspirationslängen in [s]; Prä- und Post-Phonation (SAM-Methode) ..... 145

## Tabellenverzeichnis

Tabelle 1:	Projektdatenbasis für die Gruppe der Säuglinge mit (LKGS) und ohne (KG) oro-faziale Spaltbildungen .....	18
Tabelle 2:	Deskriptive Statistik der absoluten Zeitdifferenz der Lautlängen (Lges [s]) bei Messung am Zeitsignal bzw. der Melodiekontur .....	99
Tabelle 3:	Absolute (oben) und relative (unten) Häufigkeiten der Melodiestrukturkategorien in beiden Probandengruppen.....	102
Tabelle 4:	Deskriptive Statistik der Verteilungseigenschaften der Gesamtlängen (Lges [s]) aller analysierten Vokalisationen (SAM-Methode).....	104
Tabelle 5:	Deskriptive Statistik für die Gesamtlängen (Lges [s]); SAM-Methode im Gruppenvergleich für 1B- Strukturen.....	106
Tabelle 6:	Deskriptive Statistik für die Gesamtlängen (Lges [s]); MM-Methode im Gruppenvergleich für 1B-Strukturen.....	106
Tabelle 7:	Messwerte der Vokalisationslängen (Lges [s]) mit der semi-automatischen Methode (SAM) und der manuellen Methode (MM) für Strukturen des Typs 1B (I1 und I2 = Inspirationslängen [s]; Pb = Proband).....	107
Tabelle 8:	Deskriptive Statistik für die absolute Zeitmessungsdifferenz [s] der Gesamtlänge der Vokalisationen des Typs 1B zwischen SAM- und MM-Methode beider Gruppen .....	110
Tabelle 9:	Deskriptive Statistik für die Gesamtlängen (Lges [s]); SAM-Methode im Gruppenvergleich für 2B-Strukturen.....	111
Tabelle 10:	Deskriptive Statistik für die Gesamtlängen (Lges [s]); MM-Methode im Gruppenvergleich für 2B-Strukturen.....	112
Tabelle 11:	Messwerte aller Vokalisationslängen (Lges [s]) mit der semi-automatischen Methode (SAM) und der manuellen Methode (MM) für Strukturen des Typs 2B (I1 und I2 [s] = Inspirationslängen).....	112
Tabelle 12:	Deskriptive Statistik der absoluten Zeitmessungsdifferenzen [s] für die Gesamtlängen (Lges [s]) der Vokalisationen des Typs 2B zwischen SAM- und MM-Methode beider Gruppen .....	113
Tabelle 13:	Deskriptive Statistik für die Gesamtlängen (Lges [s]); SAM-Methode im Gruppenvergleich für 3B- und MB-Strukturen .....	115
Tabelle 14:	Deskriptive Statistik für die Gesamtlängen (Lges [s]); MM-Methode im Gruppenvergleich für 3B- und MB-Strukturen ...	115

Tabelle 15:	Messwerte der Vokalisationslängen (Lges [s]) mit der semi-automatischen Methode (SAM) und der manuellen Methode (MM) für Strukturen des Typs 3B und MB (I1 und I2 [s] = Inspirationslängen, QSS = Strukturkategorie) ..... 116
Tabelle 16:	Deskriptive Statistik absolute Zeitmessungsdifferenzen [s] für die Gesamtlänge der Vokalisationen des Typs 3B/MB zwischen SAM- und MM-Methode für beide Gruppen ..... 116
Tabelle 17:	Deskriptive Statistik für die Gesamtlängen (Lges [s]); SAM-Methode im Gruppenvergleich für 1S- Strukturen ..... 118
Tabelle 18:	Deskriptive Statistik für die Gesamtlängen (Lges [s]); MM-Methode im Gruppenvergleich für 1S- Strukturen ..... 118
Tabelle 19:	Messwerte der Vokalisationslängen (Lges [s]) mit der semi-automatischen Methode (SAM) für Strukturen des Typs 1S (I1 und I2 = Inspirationslängen; B1 und B2 = Bogenlängen; p12 = Oszillationspausenlänge; alle in [s]) ..... 119
Tabelle 20:	Messwerte der Vokalisationslängen (Lges [s]) mit der manuellen Methode (MM) für Strukturen des Typs 1S (I1 und I2 = Inspirationslängen; B1 und B2 = Bogenlängen; p12 = Oszillationspausenlänge; alle in [s]) ..... 120
Tabelle 21:	Deskriptive Statistik für die Teilzeiten der Melodiekonturelemente des Strukturtyps 1S (SAM-Methode) in [s] ..... 122
Tabelle 22:	Deskriptive Statistik für die Teilzeiten der Melodiekonturelemente des Strukturtyps 1S (MM-Methode) in [s] ..... 123
Tabelle 23:	Deskriptive Statistik: Zeitmessungsdifferenzen [s] für die Gesamtlänge der Vokalisationen zwischen SAM- und MM-Methode für beide Gruppen ..... 126
Tabelle 24:	Deskriptive Statistik: Zeitmessungsdifferenzen [s] für den Melodiebogen vor der Pause (B1) der Vokalisationen zwischen SAM- und MM-Methode für beide Gruppen ..... 127
Tabelle 25:	Deskriptive Statistik: Zeitmessungsdifferenzen [s] für die Pausenlänge der Vokalisationen zwischen SAM- und MM-Methode für beide Gruppen ..... 128
Tabelle 26:	Deskriptive Statistik: Zeitmessungsdifferenzen [s] für den Melodienbogen nach der Pause (B2) der Vokalisationen zwischen SAM- und MM-Methode für beide Gruppen ..... 129
Tabelle 27:	Deskriptive Statistik für die Gesamtlängen (Lges [s]); SAM-Methode im Gruppenvergleich für 2S-Strukturen ..... 130
Tabelle 28:	Deskriptive Statistik für die Gesamtlängen (Lges [s]); MM-Methode im Gruppenvergleich für 2S-Strukturen ..... 131

Tabelle 29:	Messwerte aller Vokalisationslängen (Lges [s]) mit der semi-automatischen Methode (SAM) für Strukturen des Typs 2S (I1 und I2 = Inspirationslängen; B1-B3 = Bogenlängen; p12, p23= Oszillationspausenlängen; alle in [s]) .....	131
Tabelle 30:	Messwerte aller Vokalisationslängen (Lges [s]) mit der manuellen Methode (MM) für Strukturen des Typs 2S (I1 und I2 = Inspirationslängen; B1-B3 = Bogenlängen; p12, p23 = Oszillationspausenlängen; alle in [s]) .....	132
Tabelle 31:	Deskriptive Statistik der ermittelten Differenzen [s] zwischen SAM und MM-Methode für Strukturen der Strukturkategorie 2S .....	132
Tabelle 32:	Deskriptive Statistik: Zeitmessungsdifferenzen [s] der einzelnen Strukturelemente der LKGS-Gruppe (2S) .....	133
Tabelle 33:	Deskriptive Statistik für die Gesamtlängen (Lges [s]); SAM-Methode im Gruppenvergleich für 3S-Strukturen .....	135
Tabelle 34:	Deskriptive Statistik für die Gesamtlängen (Lges [s]); MM-Methode im Gruppenvergleich für 2S-Strukturen .....	136
Tabelle 35:	Messwerte der Vokalisationslängen (Lges [s]) mit der semi-automatischen Methode (SAM) ermittelt für Strukturen des Typs 3S (I1 und I2 = Inspirationslängen; alle in [s]) .....	136
Tabelle 36:	Messwerte der Bogenlängen und Oszillationspausenlängen mit der semi-automatischen Methode (SAM) ermittelt für Strukturen des Typs 3S (B1-B4 = Bogenlängen; p12, p23, p34 = Oszillationspausenlängen; alle in [s]) .....	137
Tabelle 37:	Messwerte der Vokalisationslängen (Lges [s]) mit der manuellen Methode (MM) ermittelt für Strukturen des Typs 3S (I1 und I2 = Inspirationslängen; B1-B4 = Bogenlängen; p12, p23, p34 = Oszillationspausenlängen; alle in [s]) .....	137
Tabelle 38:	Deskriptive Statistik für die Gesamtlängen (Lges [s]); SAM-Methode im Gruppenvergleich für KS-Strukturen .....	139
Tabelle 39:	Deskriptive Statistik für die Gesamtlängen (Lges [s]); MM-Methode im Gruppenvergleich für KS-Strukturen .....	139
Tabelle 40:	Messwerte der Vokalisationslängen (Lges [s]) mit der semi-automatischen Methode (SAM) ermittelt für Strukturen des Typs KS (I1 und I2 = Inspirationslängen; B1 und B2 = Bogenlängen; p12 = Oszillationspausenlänge; alle in [s]) .....	140



Tabelle 41:	Messwerte der Vokalisationslängen (Lges [s]) mit der manuellen Methode (MM) ermittelt für Strukturen des Typs KS (I1 und I2 = Inspirationslängen; B1 und B2 = Bogenlängen; p12 = Oszillationspausenlänge; alle in [s]).....	140
Tabelle 42:	Deskriptive Statistik: Gesamtlänge der Oszillationspausen [s] im Methodenvergleich der Strukturkategorie KS .....	141
Tabelle 43:	Deskriptive Statistik der ermittelten Differenzen [s] zwischen SAM und MM-Methode für Strukturen der Strukturkategorie KS .....	142
Tabelle 44:	Deskriptive Statistik für die Gesamtlängen (Lges [s]); SAM-Methode im Gruppenvergleich für fuzzy class-Strukturen .....	143
Tabelle 45:	Deskriptive Statistik für die Gesamtlängen (Lges [s]); MM-Methode im Gruppenvergleich für fuzzy class-Strukturen.....	144
Tabelle 46:	Messwerte der Vokalisationslängen (Lges [s]) mit der semi-automatischen Methode (SAM) und der manuellen Methode (MM) für Strukturen des Typs fuzzy class (I1 und I2 = Inspirationslängen in [s]) .....	144
Tabelle 47:	Deskriptive Statistik Inspirationslängen [s] (SAM-Methode)..	146
Tabelle 48:	Differenzen der gemessenen Vokalisationslängen Lges [s] auf Einzelsignallevel zwischen SAM und MM-Methode der Strukturkategorie 1B .....	ix
Tabelle 49:	Differenzen der gemessenen Vokalisationslängen Lges [s] auf Einzelsignallevel zwischen SAM und MM-Methode der Strukturkategorie 1S .....	xi
Tabelle 50:	Differenzen der gemessenen Vokalisationslängen Lges [s] auf Einzelsignallevel zwischen SAM und MM-Methode der Strukturkategorie 2B .....	xii
Tabelle 51:	Differenzen der gemessenen Vokalisationslängen Lges [s] auf Einzelsignallevel zwischen SAM und MM-Methode der Strukturkategorie 2S .....	xiii
Tabelle 52:	Differenzen der gemessenen Vokalisationslängen Lges [s] auf Einzelsignallevel zwischen SAM und MM-Methode der Strukturkategorie 3B/MB .....	xiii
Tabelle 53:	Differenzen der gemessenen Vokalisationslängen Lges [s] auf Einzelsignallevel zwischen SAM und MM-Methode der Strukturkategorie 3S .....	xiv
Tabelle 54:	Differenzen der gemessenen Vokalisationslängen Lges [s] auf Einzelsignallevel zwischen SAM und MM-Methode der Strukturkategorie KS .....	xiv



## Abkürzungsverzeichnis

B	Bogenlänge
CDAP	Cry Data Analysis Program
CSL	Computerized Speech Lab
e	Expiration/ expiratorisches Intervall
fo	Grundfrequenz
I	Inspirationslänge
i	Inspiration/ inspiratorisches Intervall
KG	Kontrollgruppe
Lges	Gesamtlänge
LKGS (CLP)	Lippen-Kiefer-Gaumen-Segelspalte (Cleft-Lip-Palate)
MCI	Melodiekomplexitätsindex (Melody Complexity Index)
MM	Manuelle Methode
p	Pausenlänge (innerlautliche Pause)
Pb	Proband
s/ms	Sekunde/Millisekunde
SAM	Semiautomatische Methode
SI	Segmentierungsindex
UKW	Universitätsklinikum Würzburg
ZVES	Zentrum für Vorsprachliche Entwicklung und Entwicklungsstörungen

# 1 Einleitung

## 1.1 Vorsprachliche Diagnostik

Die Vorsprachliche Diagnostik am Zentrum für vorsprachliche Entwicklung und Entwicklungsstörungen (ZVES; Leitung: Prof. Kathleen Wermke) der Poliklinik für Kieferorthopädie (Direktorin: Prof. Angelika Stellzig-Eisenhauer) am Universitätsklinikum in Würzburg befasst sich mit der Aufzeichnung, Auswertung und therapiebegleitenden Verwendung akustischer Lautäußerungen von Säuglingen und Kleinkindern mit einem Risiko für eine Sprech- und/oder Sprachentwicklungsstörung. Bereits seit den 1960er Jahren gibt es Bestrebungen verschiedener Arbeitsgruppen, Informationen aus Babylauten als entwicklungsdiagnostische Zusatzinformationen zu verwenden (Barr, Hopkins, & Green, 2000; Wasz-Hockert, Lind, Partanen, Valanne, & Vuorenkoski, 1968; Wasz-Höckert, Michelsson, & Lind, 1985). Lange bevor der Säugling damit beginnt erste Wörter zu bilden setzt die Sprachentwicklung ein. Bereits im Mutterleib nimmt das Ungeborene unterschiedlichste akustische Phänomene in seiner Umgebung wahr und beginnt diese zu verarbeiten (Byers-Heinlein, Burns, & Werker, 2010; DeCasper & Spence, 1986; Kisilevsky et al., 2003). Vom Weinen über Cooin, einfachem Babbeln, kanonischem Babbeln und schließlich ersten Worten macht der Säugling bereits in den ersten 12 Monaten eine bemerkenswerte Entwicklung auf dem Weg zur Sprache durch. Die Lautäußerungen des Neugeborenen und jungen Säuglings verfügen über eine höhere Komplexität und einen dichteren Informationsgehalt als lange Zeit und teilweise noch heute in der Fachliteratur angenommen. Säuglingsweinen geht weit über reine Schmerz- und Unmutsbekundungen hinaus und repräsentiert durch ein angeborenes Melodieentwicklungsprogramm (Wermke & Mende, 2011/2012, 2016) bereits eine entscheidende Phase der vorsprachlichen Entwicklung.

Wie das Sprechen basieren auch die frühen Stadien der Lautbildung auf einem komplexen Zusammenspiel muskulärer und neurophysiologischer Vorgänge. Wermke weist bereits 1987 darauf hin, dass die Kehlkopfmuskeln zu den schnellsten muskulären Effektoren des menschlichen Gehirns gehören und

bereits geringfügige Störungen des neurophysiologischen Regelnetzwerkes unter Umständen in einer ausgeprägten Veränderung vorsprachlicher Lauteigenschaften resultieren (Wermke, 1987). Daher dient die am ZVES entwickelte *Vorsprachliche Diagnostik* als nicht-invasives, die medizinische Diagnostik begleitendes Instrumentarium, um nicht nur den aktuellen Entwicklungsstand des Säuglings einschätzen zu können, sondern auch nach Frühindikatoren für potenzielle hormonell, neurologisch und muskulär bedingte (Sprach-) Entwicklungsstörungen bei Risikokindern zu suchen. Diese Störungen sind sonst häufig erst im Sprachalter deutlich zu erkennen. Eine Früherkennung ermöglicht es, eine Therapie möglicherweise bereits im Säuglingsalter einzuleiten.

Aber gerade Patientenkollektive, die im Rahmen der Vorsprachlichen Diagnostik besonders von einer Frühdiagnose profitieren würden, stellen sich als überaus heterogen dar: dazu zählen Säuglinge mit angeborenen Lippen- Kiefer- Gaumen und Segelspalten, mit hochgradig sensorineuralen Hörstörungen, Autismus-Spektrum-Störungen und Frühgeborene. Auch innerhalb der genannten Patientengruppen variiert die individuelle vorsprachliche Entwicklung der Säuglinge trotz gleicher Diagnose und vergleichbarer medizinischer sowie sozioemotionaler, familiärer Faktoren zum Teil erheblich.

Besonders deutlich wird die hohe Heterogenität und Variabilität bei der Patientengruppe von Säuglingen mit angeborenen Lippen- Kiefer- Gaumen- und Segelspalten (LKGS) bereits bei Betrachtung der klinisch gebräuchlichen LAHSHAL-Klassifikation (Koch, Koch, Grzonka, & Gundlach, 2003)<sup>1</sup>. Bedingt durch die zeitlich unterschiedliche embryonale Entwicklung von Lippen und Gaumen (Schwenzer & Bacher, 2011) entstehen unterschiedliche Ausprägungen, die unter derselben Entität LKGS zusammengefasst werden. Die Bandbreite der unterschiedlichen Konditionen umfasst eine Vielfalt von

---

<sup>1</sup> Beschreibt alle betroffenen Regionen an Oberlippe (L), Alveolarfortsatz des Oberkiefers (A), Hartgaumen (H), Segelgaumen (S), evtl. auch äußerer Nase (N) und Vomer (V) (vgl. Koch et al. (2003))

Diagnosen, die unter Anderem von einseitigen submukösen Lippenspalten bis zu beidseitigen kompletten LKGS-Spalten reichen. Diese werden oft auch zusätzlich begleitet von bekannten oder verdeckten Syndromen. Zusätzlich erschweren die häufig in Verbindung mit Spaltenbildung auftretenden Paukenergüsse und die daraus resultierenden Schalleitungsstörungen das Sprachverständnis und damit den Spracherwerb (Schönweiler, Schönweiler, Schmelzeisen, & Ptok, 1995; Wendler, 2015). Bei Säuglingen mit LKGS-Spalten sind bereits die Lauteigenschaften frühester vorsprachlicher Vokalisationen im Vergleich zu Nichtspaltträgern verändert (Wermke et al., 2011; Wermke, Hauser, Komposch, & Stellzig, 2002). Die Varianten führen zu unterschiedlich starken Beeinträchtigungen und Variationen der Lautbildung und Sprachentwicklung (Eiband; Grunwell, 1993; Howard, 2011; Lederman, Zmora, Hauschildt, Stellzig-Eisenhauer, & Wermke, 2008; Wermke et al., 2002; Wermke et al., 2011). Dadurch entsteht eine hohe inter- und intraindividuelle Variabilität, deren Beherrschung und klinische Nutzbarmachung geeignete Entwicklungsmodelle und die möglichst automatische Analyse großer Stichproben und Datenmengen erfordert. Am ZVES wird dazu geforscht (z.B. Birr, 2009; Dirauf, 2010; Fuamenya, 2011; Ostermann, 2010; Pelzer, 2011; Sereschk, 2019; Wermke, Quast, & Hesse, 2018).

Obwohl sich seit Beginn der Säuglingsschreiforschung in den 1960er Jahren die technischen Methoden zur Analyse auditiver Signale rasant weiterentwickelt haben und immer komplexere und leistungsfähigere Verfahren entwickelt wurden, die inzwischen auch den Alltag stark prägen (Sprachsteuerung für Smartphones, Smart-Home etc.), sind Anwendungen im Bereich der vorsprachlichen Sprech- und Sprachdiagnostik von Kindern kaum vorhanden. Ein wesentlicher Grund hierfür ist die hohe Variabilität der akustischen Merkmale in Säuglings- und Kinderlauten bei gleichzeitig fehlenden Normwerten. Selbst moderne ingenieurtechnische Methoden zur Gewinnung medizinisch-biologischer Informationen aus dem Säuglingsschrei vernachlässigen fast ausnahmslos wichtige Einflussfaktoren wie Gestationsalter (Shinya, Kawai, Niwa, & Myowa-Yamakoshi, 2014), Geburtsmodus (Höing, 2010), Alter (Wermke, 2002), oder Geschlecht (Wermke, Hain, Oehler, Wermke, & Hesse, 2014;

Wermke, Quast et al., 2018). Auch andere Faktoren, wie endokrine Bedingungen (Minipubertät) und die Umgebungssprache, deren Einfluss in unterschiedlichen Publikationen nachgewiesen wurden (Borysiak et al., 2017; Mampe, Friederici, Christophe, & Wermke, 2009; Shinya et al., 2014; Wermke et al., 2014; Wermke et al., 2016), werden in bisher publizierten technischen Ansätzen zur automatisierten vorsprachlichen Lautanalyse noch vernachlässigt. Auch die oft ersatzweise Zugrundelegung des Lautproduktionsmodelles der Erwachsenensprache bei der Analyse von Säuglingslauten ist mit methodischen Limitationen verbunden. Das kindliche Lautproduktionssystem befindet sich unter einer kontinuierlichen Reifung und Lauteigenschaften weisen neben einer hohen altersabhängigen Variabilität und reifungsbedingten Veränderung Spezifika auf. Keiner der bisherigen Ansätze zur automatisierten Analyse von Säuglingslauten aus medizinischer Perspektive hat die inzwischen bekannten Entwicklungsprozesse und -phänomene berücksichtigt.

Damit die Vorsprachliche Diagnostik klinisch relevant und praktikabel wird, ist es u.a. erforderlich, die bisher aus Gründen der Zuverlässigkeit und Präzision vorwiegend manuellen oder semi-automatischen Verfahren weiter zu automatisieren. Um eine klinisch zuverlässige Unterscheidung von pathologischen und physiologischen Eigenschaften von Lautäußerungen beziehungsweise den sie charakterisierenden Messgrößen vornehmen zu können, bedarf es geeigneter, altersspezifischer Referenzwerte gesunder Kinder. Um dabei der Vielzahl an individuellen Entwicklungskonditionen und Einflussfaktoren gerecht zu werden, ist die Verarbeitung großer Datenmengen zwingend notwendig. Diese gestaltet sich durch manuelle Auswertung von Einzellauten zeitlich aufwendig und personalintensiv.

Nachdem die „Säuglingsschreidiagnostik“ in den 1960er-70er Jahren durch viele Forschergruppen in der Welt immer mehr in den Fokus des pädiatrischen Interesses gerückt war, stand man vor der Aufgabe mit den sich rapide entwickelnden Computersystemen auch nach Verfahren zu suchen, die eine automatische, objektive Analyse der Säuglingsschreie im klinischen Kontext ermöglichen. Eine Landmarke in dieser Entwicklung war die Dissertation von Howard L. Golub am Massachusetts Institute of Technology (1980).

Golub entwickelte ein Modell für die Analyse von Weinauten, das Parameter, die die laryngeale Quelle der Lautproduktion reflektieren mit solchen, die die Vokaltraktfilterfunktionen charakterisiert, kombiniert (laryngeale und supralaryngeale Merkmale). Außerdem bezog er respiratorische Kenngrößen mit ein. Jeder Einzellaut (Schrei) wurde durch ein Set von 88 Parametern charakterisiert und anschließend computergestützt analysiert. Er postulierte, dass er mithilfe dieses Systems eine Anzahl medizinischer Auffälligkeiten aus dem Säuglingsschrei detektieren und insbesondere Risikokinder für den plötzlichen Kindstod in einem Screening identifizieren könne. Eine verlässliche Einschätzung der Spezifität und Sensitivität des Verfahrens im Falle von auffälligen Säuglingen und Schreien sowie die signalanalytische Zuverlässigkeit der Verfahren kann nicht beurteilt werden. Die Arbeit selbst und die vorhandene Fachliteratur liefert keine geeignete Information, mit der eine solche Einschätzung möglich wäre (Golub & Corwin, 1982). Die Tatsache, dass auch 40 Jahre später noch kein marktreifes System vorhanden ist, ist zumindest ein starkes Indiz dafür, dass die formulierten verheißungsvollen Ergebnisse in nachfolgenden Studien nicht bestätigt werden konnten. Die Arbeit Golubs hat jedoch Pioniercharakter und gilt als erster Ansatz für eine systematische und objektive Schreianalyse (Barr et al., 2000). Abou-Abbas et. al (2017) berichten über verschiedene weitere frühere Klassifikationssysteme, mit denen unterschiedliche Autoren pathologische oder emotionale Zustände von Säuglingen automatisch erkennen wollten. Die berichteten Trefferquoten an den Testsignalen sind hoch: mit einer Trefferquote zwischen 88-100% sollen beispielsweise „hypoakustische“ Säuglinge von gesunden Säuglingen anhand der Säuglingsschreie differenzierbar sein (Hariharan, Sindhu, & Yaacob, 2012; Reyes-Galaviz, Cano-Ortiz, & Reyes-García, 2008; Sanfeliu, Ruiz-Shulcloper, Orozco-García, & Reyes-García, 2003), und mit einer Trefferquote von 93,16% bzw. 94% sollen Säuglinge mit „Asphyxie“ von gesunden unterscheidbar sein (Sahak, Lee, Mansor, Yassin, & Zabidi, 2010; Zabidi, Khuan, Mansor, Yassin, & Sahak, 2010). Angst-, Hunger- und Schmerzdetektion sollen mit einer Trefferquote von 90,4% aus Daten der ersten Sekunde von aufgezeichneten



Schreien nach Angst-, Hunger- oder Schmerzreiz möglich sein (Petroni, Malowany, Johnston, & Stevens, 1995).

Die publizierten hohen Erkennungsraten sind möglicherweise für das jeweils verwendete Testsignalsetting erreicht worden, sind jedoch nicht generalisierbar – ein Umstand, den die Autoren oft nicht mitberichten. Angesichts der Tatsache, dass auch von diesen Ansätzen tatsächlich keiner bisher klinisch relevant wurde, sind weitere gezielte Untersuchungen notwendig, die in enger Zusammenarbeit von spezialisierten Informatikern mit Ärzten erfolgen müssen. Diese Zusammenarbeit ist notwendig, um klinisch relevante Fragestellungen für die Lautanalyse und Vorsprachliche Diagnostik zu ermitteln und die notwendigen und vor allem medizinisch und entwicklungsdiagnostisch zuverlässigen Begleitinformationen für die Audiodaten zusammenzustellen und bei der Klassifikation der Lautäußerungen zu berücksichtigen.

Der Weg zu einer zeitsparenden, therapiebegleitenden Vorsprachlichen Diagnostik ist jedoch aus mehreren Gründen ein schwierigeres Unterfangen als man zunächst annehmen könnte:

Der Lautbildung des jungen Säuglings liegt ein teilweise noch unreifes System (Vokaltrakt) zugrunde (Bosma, Truby, & Lind, 1965) und das laryngeale Regelsystem arbeitet aufgrund der winzigen Stimmlippen und des hohen subglottischen Drucks (Gaultier, 1995) bereits an seiner Leistungsgrenze (Wermke et al., 2019). Daher sind Signaleigenschaften der Babylaute im Vergleich zu denen der späteren Sprachlaute deutlich variabler und schwerer standardisiert zu untersuchen. Säuglingslaute (Weinen) sind aufgrund eines vergleichsweise hohen subglottischen Drucks unter anderem häufig geprägt von nichtlinearen Turbulenzen (stimmlose Elemente = phonatorische Rauschbanden), plötzlichen Frequenzsprüngen (shift) und Subharmonischen (Mende, Herzel, & Wermke, 1990). Diese Lauteigenschaften in den von Säuglingen aufgezeichneten Sequenzen (Audiofiles) stellen hohe Anforderungen an eine automatische Segmentierung der interessierenden Laute und deren Analyse durch technische Systeme. Im Unterschied zur Sprachanalyse ist vor allem die zuverlässige Identifizierung der interessierenden Abschnitte aus einer

aufgezeichneten Audiosequenz eine Herausforderung für die automatische Analyse. Eine Sequenz aufgezeichneter Weinlaute (Langserie) besteht in der Regel aus einer Abfolge von expiratorischen Weinlauten unterschiedlicher Intensität und Dauer (Einzellaute), Inspirationsgeräuschen, Pausen und oft auch Hintergrundgeräuschen, wie z.B. überlagertes Sprechen einer Kontaktperson und/oder Geräusche des klinischen Alltagsbetriebs. Eine Aufzeichnung von Säuglingslauten in der Schallkabine des ZVES lässt sich aus organisatorischen und logistischen Gründen nur in Ausnahmefällen realisieren. Eine semi- oder zukünftig vollautomatische Identifizierung (und Segmentierung) relevanter Zeitabschnitte (Laute) und anschließende Analyse ist daher ein notwendiges Desideratum der Vorsprachlichen Diagnostik. In Abbildung 1 ist ein typisches Beispiel einer solchen Langserie zu sehen: Die Abbildung zeigt im oberen Drittel eine Zeitbereichsdarstellung (0-13s) und in den unteren zwei Dritteln ein Schmalbandspektrogramm (linear bis 4kHz) der enthaltenen Einzellaute und ingressiven Geräusche. Ein Einzellaut ist hier als lautliche Expression während einer Expiration definiert (Robb, Sinton-White, & Kaipa, 2011; Wermke & Robb, 2010). Die Abbildung verdeutlicht die Variabilität in der Zeitdauer der Einzellaute, die sich mit vergleichsweise kurzen, aber ebenfalls variablen Inspirationspausen abwechseln. Diese Pausen enthalten teilweise inspiratorische Phonationen, wie sie in den ersten 10 Lebenswochen regelhaft bei gesunden Säuglingen beobachtet werden (Wermke, Haschemi, Hesse, & Robb, 2018). Diese müssen von den expiratorischen Phonationen unterschieden werden, da ihnen unterschiedliche respiratorisch-laryngeale Kontrollmechaniken zugrunde liegen (Robb et al., 2019; Sereschk, 2019; Wermke, Haschemi et al., 2018).

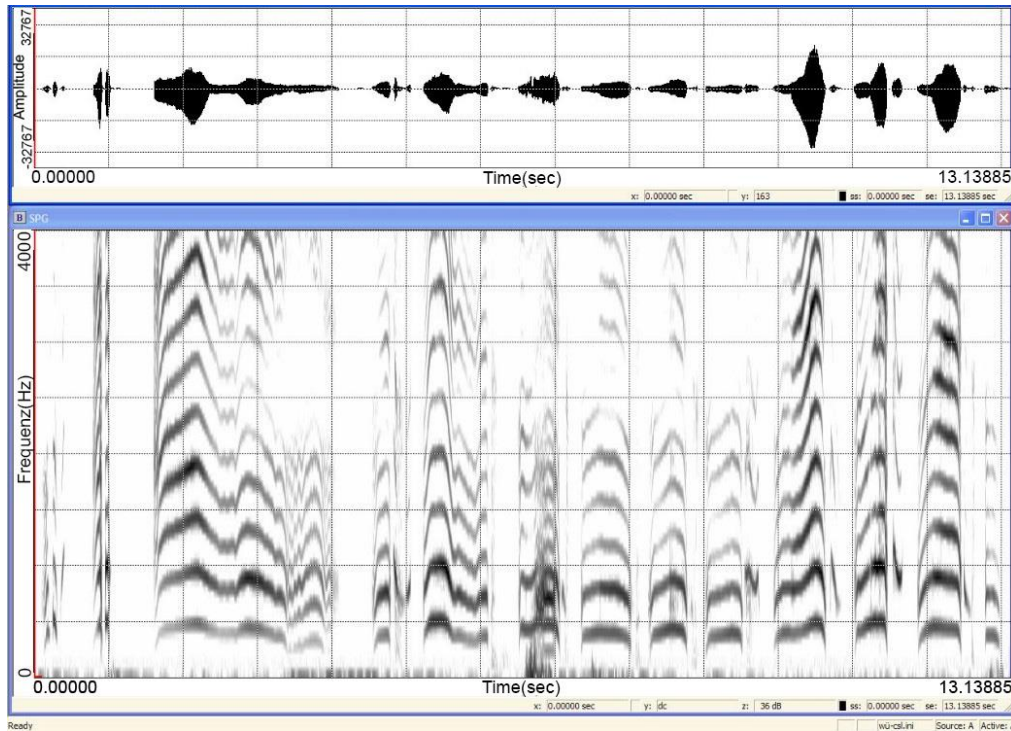


Abbildung 1: Zeitbereichsdarstellung und Schmalbandspektrum bis 4 kHz (y-Achse) 0-13s (x-Achse) einer Abfolge von Weinlauten und Inspirationsgeräuschen [KAY-CSL (KAY Pentax Medical)]

Um die jeweils diagnostisch relevanten expiratorischen Abschnitte (Einzellaute) für die weitere Analyse verfügbar zu machen, muss zunächst eine Annotation aller unterschiedlichen in der Langserie enthaltenen Phonationen stattfinden, um sie nachfolgend automatisch selektieren zu können (Segmentierung).

Diese automatische Segmentierung ist daher auch Gegenstand aktueller Ansätze zur Nutzbarmachung der Vorsprachlichen Diagnostik aus medizinischer Perspektive. Die derzeit führende Gruppe auf diesem Gebiet (am Department of Electrical Engineering der École de Technologie Supérieure (Quebec)) hat sich das Ziel der Entwicklung eines vollautomatischen, non-invasiven Systems zur Diagnose von neonatalen Erkrankungen gesetzt, das allein auf der Schreianalyse basiert: „Specifically, this is the main goal of our project: to develop a fully automatic noninvasive system able to diagnose diseases based solely on cry sound analysis.“ (Abou-Abbas et al., 2017, p. 1318). Sie will damit einen Beitrag zur Senkung der Säuglingssterblichkeit in den ersten Lebenswochen leisten (ibid.). Um ein solches System zu entwickeln, hat diese kanadische Arbeitsgruppe als ersten Schritt ebenfalls die Notwendigkeit einer Segmentierung

der im klinischen Kontext aufgenommenen Lautäußerungen erkannt und in Angriff genommen. Der Hauptfokus der Forschung dieser Gruppe liegt daher bisher auf der vollautomatischen Segmentierung von aufgezeichneten Sequenzen der Säuglingsvokalisationen (Weinen). Die Autoren experimentieren bei der Segmentierung unter anderem mit den in der Spracherkennung populären stochastischen Suchmethoden wie Hidden Markov Models und Gaussian Mixture Models (Abou-Abbas et al., 2017, p. 1322) und vergleichen die Ergebnisse, um eine möglichst korrekte Annotation der in den Langserien enthaltenen Laute, Inspirationsphonationen beziehungsweise Hintergrundgeräusche zu erreichen. Die Gruppe gibt für ihre Testserien eine Erfolgsrate von bis zu 94,29% an (Abou-Abbas et al., 2017, p. 1329). Bisher ist das Verfahren aber noch in der Testphase und steht nicht zur Verfügung.

Die Säuglingslautanalyse am ZVES verfolgt neben der ebenfalls im Fokus stehenden Segmentierungsproblematik vorerst noch einen anderen Ansatz als die Entwicklung einer „Standalone“-Lösung für den klinischen Alltag (Abou-Abbas et al., 2017). Durch die große Variabilität der respiratorisch-laryngealen Reife- und Funktionszustände des im Fokus stehenden Patientenkollektivs und immer noch unzureichenden Referenzdaten erscheinen Bestrebungen eines „simplen“ vollautomatischen Klassifikationsansatzes für „normale“ und „pathologische“ Säuglingslaute momentan noch verfrüht. Die systematische, in der Forschung angewandte semi-automatische Analyse in einem interdisziplinären Setting mit Kieferorthopäden, Neonatologen, Pädiatern, HNO-Ärzten, Endokrinologen, Neurochirurgen und Ingenieuren erscheint derzeit geeigneter.

Die Entwicklung klinisch relevanter automatischer Auswertungssysteme muss interdisziplinär unter intensiver Mitarbeit von Ärzten und, im Falle der Patienten mit Lippen-, Kiefer-, Gaumen-, Segel-Spaltbildungen (LKGS), auch von Fachzahnärzten (Kieferorthopäden) erfolgen. Der Autor ist Zahnarzt und gleichzeitig ausreichend spezialisiert, um neben der Auswertung der spezifischen Studiendaten (Kapitel 2.1) hier auch die für die Erfüllung der Zielsetzung der Arbeit erforderlichen programmiertechnischen Aufgaben zu übernehmen.

## 1.2 Zielsetzung und Methodologie

### 1.2.1 Zielsetzung

Zielsetzung der vorliegenden Arbeit war die semi-automatische Analyse von Säuglingslauten im Rahmen eines neurobiologisch-kieferorthopädischen Forschungsprojektes. Bei dem Forschungsprojekt, aus dem die hier analysierten Daten stammten, handelte es sich um eine explorative Pilotstudie mit dem Titel „Evaluation of the Neurobiology of Neonatal Cry Patterns in Infants with Cleft Lip and/or Palate Prior to Surgery: A Pilot Study“. Das Projekt wurde durch die amerikanische Cleft Palate Foundation mit einem Grant an die Projektleiterin Amy Conrad (PhD) gefördert und am Iowa College of Medicine der Universität Iowa (USA), The Stead Family Department of Pediatrics, durchgeführt (Ethikvotum der Universität Iowa und schriftliche Einwilligung der Eltern lagen vor). Die Analyse der dort erhobenen Babylaute nach dem ZVES-Protokoll sollte am ZVES in Würzburg erfolgen. Im Rahmen vorliegender Arbeit sollten die temporalen Eigenschaften der aufgezeichneten Säuglingslaute nicht nur manuell, sondern auch semi-automatisch analysiert werden, um eine geeignete Methode für das Folgeprojekt an 200 Probanden bereitzustellen. Das Projekt ist ab 2021 geplant. Dazu sollte zunächst eine spezifische Programmroutine für das Phonetikprogramm PRAAT (Boersma & Weenink, 2018a), nachfolgend „Skript“ genannt, vom Autor erarbeitet, anhand vorgegebener Testsignale getestet und dieses Vorgehen ausführlich beschrieben werden. Die ausführliche Beschreibung sollte zukünftigen Nutzern im Rahmen der Folgestudie sowie anderer klinischer Studien ein Verständnis für das Programm und damit eine korrekte Anwendung ermöglichen. Auch nicht mit der programmeigenen Skriptsprache vertrauten Anwendern sollte es, nach Implementation des Skriptes im Audiolabor des ZVES, möglich sein, (teil-) automatisiert beliebige Langserien von vorsprachlichen Lautäußerungen im Rahmen der Vorsprachlichen Diagnostik zu annotieren, beliebige interessierende Einzelsignale herauszufiltern und die prosodischen Endparameter Melodie- und Intensitätskontur zu berechnen. Die Analyseergebnisse sollten so abgespeichert werden, dass sie anschließend an bereits installierte Systeme am ZVES, wie das Cry-Data-

Analysis-Programm (CDAP) oder Statistiksoftware wie Excel oder SPSS, übergeben werden können.

Vor Anwendungen des Skriptes auf die Forschungsprojektdaten im Kooperationsprojekt sollte dessen einwandfreie Funktion an vorgegebenen geeigneten Testsignalen verschiedener Patientengruppen belegt werden. Parallel sollten Betaversionen des Skriptes kontinuierlich im Rahmen eines anderen laufenden Projektes „Analyse von Komfortlauten (babbling) bei Säuglingen zum Zeitpunkt des Neugeborenen-Hörscreening-2“ (Ethikvotumnummer 308/17me) am ZVES getestet werden und die Rückkopplung durch die Anwender bei der Weiterentwicklung berücksichtigt werden. Die Release Candidates der Skripte sollen auch im Rahmen von Kooperationsprojekten des ZVES mit Forschergruppen in den USA, China und Neuseeland angewendet und später zur freien Verfügung für die medizinische Säuglingsforschung online gestellt werden.

Warum schloss die Zielsetzung der quantitativen Analyse temporaler Eigenschaften von Säuglingslauten im Rahmen des genannten Projektes auch einen Aspekt der Methodenentwicklung in Form eines spezifischen PRAAT-Skriptes mit ein?

Um eine korrekte Segmentierung der unterschiedlichen in den Langserien enthaltenen Komponenten (Abbildung 1) zu garantieren, erfolgt die Annotation und Segmentierung bisher vor allem manuell (z.B. Barr et al., 2000, pp. 137–156; Robb et al., 2011; Shinya et al., 2014; Wermke, Haschemi et al., 2018). Realisiert wurde dies durch manuelle Cursorsetzung unter Nutzung von Soft- und Hardwarelösungen zur Sprachanalyse, wie z.B. dem lange Zeit als Quasistandard geltenden Computerized Speech Lab (KAY Pentax Medical) oder dem inzwischen verbreiteteren Softwaresystem PRAAT (Boersma & Weenink, 2018a), mit denen (neben einer Vielzahl von anderen Funktionen) Zeitbereichsdarstellungen und Frequenzspektrogramme grafisch dargestellt werden können. In Abbildung 2 ist ein Beispiel einer Langserie mit Zeitbereichsdarstellung (oberes Drittel), zugehörigem Frequenzspektrogramm

(Mitte) sowie der manuell vorgenommenen Annotation der Langserie (unteres Drittel) mittels PRAAT zu sehen.

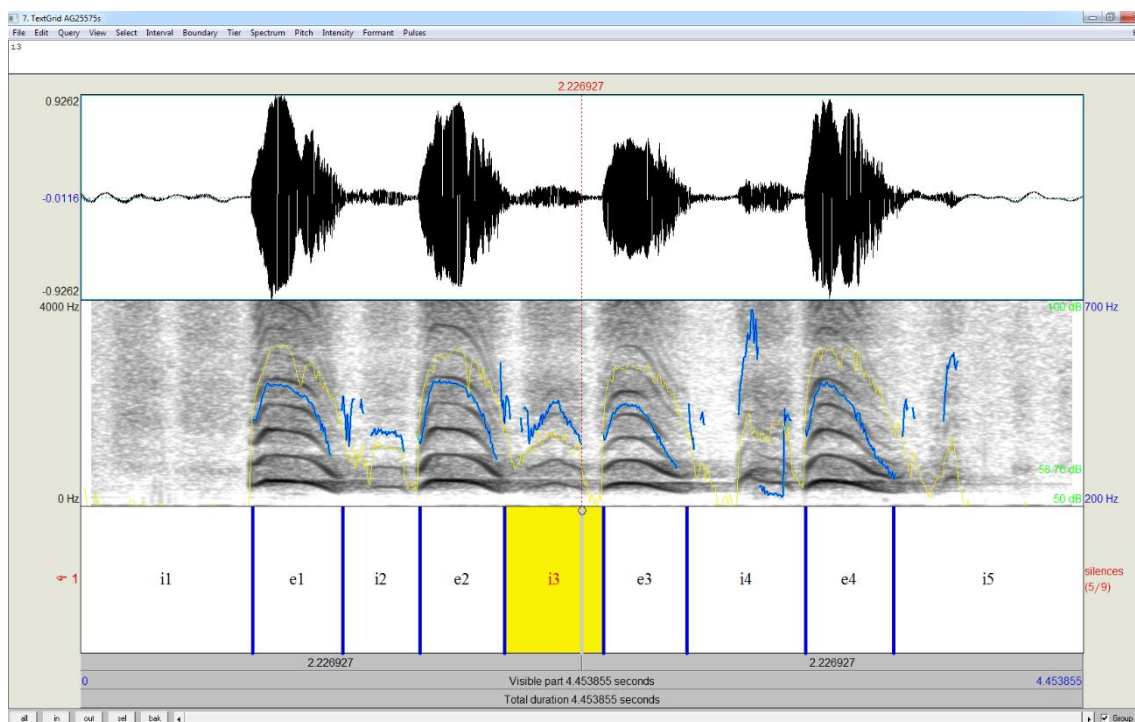


Abbildung 2: Zeitbereichsdarstellung, Schmalbandspektrogramm bis 4 kHz und Aufteilung einer Langserie von 4,45s [PRAAT (Boersma & Weenink, 2018a)]

Die Lautgrenzen wurden unter Zuhilfenahme von Zeitbereichsdarstellung und Schmalbandspektrogramm visuell aufgesucht und durch manuelle Cursorsetzung inspiratorische (i) und expiratorische (e) Einzelsignale markiert, die nach einer auditiven Kontrolle als einzelne Sounddateien anonymisiert abgespeichert werden. Die so entstandenen kürzeren und spezifisch selektierten Sounddateien wurden bisher so den spezifischen Analysen zugeführt. Der Kontext, aus dem die Einzelsignale in der Originallangserie stammten, ging dabei verloren.

Eine Analyse unter rhythmischen Gesichtspunkten in verschiedenen Zeitskalen konnte mit Lautäußerungen, die als Reaktion auf Lautäußerungen einer Kontaktperson entstanden waren, nicht erfolgen. Dies ist für ein leistungsfähiges System aber ebenso notwendig wie die zuverlässige automatische Identifikation von egressiven und ingressiven inspiratorischen Phonationen in Säuglingslauten. Inspiratorische/ ingressive Phonationen können in der Vorsprachlichen

Diagnostik bei bestimmten Syndromen und respiratorischen Dysfunktionen von Relevanz sein (Sereschk, 2019; Wermke et al., 2019; Wermke, Haschemi et al., 2018). Dies erfordert eine entsprechende Annotation und Segmentierung der Langserien, die eine Analyse ohne Verlust des Kontextes der umgebenden akustischen Ereignisse des jeweiligen extrahierten Elements erlaubt. Dieser Ansatz war bisher am ZVES noch nicht geeignet realisiert und wurde bisher auch in keiner anderen klinischen Arbeitsgruppe verfolgt. Für die Zielsetzung der vorliegenden Arbeit im Rahmen der Forschungsprojekte war gefordert, die temporalen Daten mit hoher Messgenauigkeit und unter Reproduzierbarkeit des zeitlichen Kontextes der Einzelsignale in den Originallangserien zu analysieren. Die bis dato verwendete ausschließlich manuelle Methode der Vermessung zeitlicher Eigenschaften an der Melodie war dazu nicht geeignet. Der Grund dafür sind Vorbefunde aus anderen Dissertationen, die gezeigt hatten, dass diese Laute häufig innermelodische Pausen aufweisen und die teilweise kurzen melodischen Bereiche durch laryngeale Konstriktionen stark verrauscht sind. Damit gibt es in diesen Bereichen keine Grundfrequenzwerte und die Vermessung wird fehlerhaft. Eine Vermessung anhand des Zeitsignals war daher dringend erforderlich und wurde in der vorliegenden Arbeit gleich semi-automatisch realisiert.

Zusammengefasst hatte die vorliegende Arbeit folgende Zielsetzung:

- Semi-automatische Segmentierung und quantitative Analyse der Projektdaten
- Entwicklung, Testung (anhand vorgegebener Testsignale aus dem ZVES-Archiv) und Implementierung hierzu geeigneter PRAAT-Skripte, auch für Folgeprojekte mit erheblich größerer Probandenzahl
- Anwendung der Skripte auf die Projektdaten und Vergleich der Ergebnisse mit bisher gebräuchlichen (manuellen) Analyseroutinen am selben Datensatz im ZVES-Signalanalyselabor
- Finale Zusammenstellung der Analyseergebnisse der Projektdaten für Publikationen



## 1.2.2 Methodologie

Die Vorgehensweise zur Erfüllung der in Kapitel 1.2.1 dargestellten Zielsetzung vorliegender Arbeit bestand zunächst in einer Voranalyse und Sichtung der pseudonymisierten Originallangserien (Audiofiles) von amerikanischen Säuglingen mit und ohne oro-fasziale Spaltbildungen aus genanntem Forschungsprojekt (siehe Kapitel 2).

Dabei zeigte sich, dass die vorliegenden Rohdaten notwendige Qualitätsmerkmale erfüllten: Sie wurden individuell und standardisiert von geschultem Personal aufgenommen und angesteuert sowie in einer lärmreduzierten Umgebung aufgezeichnet, um einen möglichst hohen Signal-Rausch-Abstand (ca. 40 dB) zu erreichen.

In einem nächsten Schritt wurden die zu analysierenden Messgrößen festgelegt: Für die temporale Analyse sollten die Gesamtlänge der Einzelsignale (Laute während einer Expirationsphase) sowie zusätzlich, bei durch laryngeale Oszillationspausen rhythmisierten Lauten, Partiallängen der einzelnen Lautanteile sowie der Pausen dazwischen bestimmt werden. Gerade letzteres war wesentlich, da solche durch innermelodische Pausen segmentierte Laute bei Säuglingen mit oro-faszialen Spaltbildungen besonders häufig vorkommen (Wermke et al., 2011; Ostermann, 2010). Die Pilotstudie, für die die Analysen erfolgten, hatte einen Vergleich zwischen morphologischen Hirnstrukturen und Lauteigenschaften bei Säuglingen mit und ohne oro-fasziale Spaltbildungen zum Ziel (Kapitel 2.1). Um erstmalig auch rhythmische Eigenschaften in der Originalzeitsequenz berücksichtigen zu können, musste die zeitliche Abfolge der Einzellaute erhalten bzw. eindeutig reproduzierbar sein (vgl. Kap.1.2.1).

Ein dazu geeignetes Skript für das zu verwendende Analyseprogramm (PRAAT) lag am ZVES nicht vor und musste daher vom Autor vorliegender Arbeit zur Erfüllung der Zielsetzung erarbeitet werden:

Bei PRAAT handelt es sich um ein quelloffenes Phonetik- und Linguistikprogramm, das von den niederländischen Phonetikern Paul Boersma und David Weenink der Universität von Amsterdam entwickelt und kontinuierlich

erweitert wurde (Boersma & Weenink, 2018a). Es kann laut der Autoren dazu eingesetzt werden, Sprache zu „analysieren (Grundton-, Intensitäts-, Formantenanalyse und weitere), zu synthetisieren und zu manipulieren“ [übersetzt aus dem Englischen (Boersma & Weenink, 2018b)]. Darüber hinaus können erzeugte Analyseergebnisse grafisch dargestellt, in Tabellen gespeichert oder statistisch analysiert werden. Es wird neben seiner originären Bestimmung auch seit längerer Zeit schon erfolgreich in der Analyse vorsprachlicher Lautäußerungen eingesetzt (z.B. Clement, 2004; Esteve-Gibert & Prieto, 2013; Wermke et al., 2013). Das quelloffene Design erlaubt es, auch eigene Versionen zu kompilieren. Auch die kontinuierliche Weiterentwicklung, gute Updatepolitik der Entwickler und nicht zuletzt der breite Funktionsumfang tragen zur Beliebtheit und großen Verbreitung bei. PRAAT hat inzwischen, fast weltweit, das CSL-System (KAY Pentax Medical) als Quasistandard in der bioakustischen und linguistischen Forschung und Praxis abgelöst. Am ZVES sind beide Systeme neben weiteren speziellen Analyseprogrammen vorhanden und werden kombiniert für die Vorsprachliche Diagnostik eingesetzt. Die in Kapiteln 2.4- 2.4.3 und 3.1 sowie Anhang A 1 formulierten Beschreibungen sollen dazu dienen, wichtige Funktionen und Programmumgebungen von PRAAT sowie den Quellcode der Skripte zu erläutern, um auch dem nicht mit der Software vertrauten Leser und zukünftigen Nutzern der Programme das Verständnis der in der vorliegenden Arbeit erarbeiteten und beschriebenen Programmroutine zu erleichtern. Für eine ausführlichere Beschreibung wird auch auf die offizielle Anleitung der PRAAT- Autoren verwiesen, auf die sich auch die in den Kapiteln 2.4- 2.4.3 gegebene Beschreibung stützt (Boersma & Weenink, 2018b).

Die Herausforderung bei der hier gestellten Aufgabe der semi-automatischen Analyse temporaler Lauteigenschaften von Säuglingen mit oro-faszialen Spaltbildungen bestand in dem hohen Anteil segmentierter Melodien und der Differenzierung dieser innermelodischen Pausen von Pausen aufeinanderfolgender Einzellaute. In der Anwendung der auf das konkrete Forschungsprojekt vorangehender Skriptentwicklungs- und Testphase wurde dies berücksichtigt. In einer audio-visuellen Voranalyse wurden die Testsignale charakterisiert und die erstellten Schmalbandspektrogramme grafisch dargestellt

und beschrieben (Kap. 2.2.). Die in Kapitel 2.2 beschriebenen typischen Lauteigenschaften sollten bei der Skriptentwicklung berücksichtigt werden. Gleichzeitig sollte die Anwendung des Skriptes auf die Testsignale die prinzipielle Eignung der zu entwickelnden semi-automatischen Programmroutine prüfen, bevor deren Anwendung auf die Projektdaten erfolgen konnte.

Die Datenanalyse erfolgte durch eine deskriptive Statistik der ermittelten temporalen Daten und vergleichenden Testung zwischen beiden Probandengruppen der Pilotstudie (Kapitel 2.1).

## 2 Material und Methoden

Entsprechend der Zielsetzung der vorliegenden Arbeit erforderte die Aufgabe der semi-automatischen Analyse temporaler Lauteigenschaften unter Erhalt der zeitlichen Kontextinformation (siehe Kapitel 1.2.2) auch eine spezifische Skriptentwicklung für das Phonetikprogramm PRAAT (Boersma & Weenink, 2018a), inklusive seiner ausführlichen Dokumentation und Testung, um das geplante Folgeprojekt an einer erheblich größeren Stichprobe vorzubereiten. Die Testung erfolgte an vorgegebenen Testsignalen aus dem Archiv am ZVES. Nach erfolgreicher Testung sollte das Skript im Rahmen einer Datenanalyse für das Forschungsprojekt „Evaluation of the Neurobiology of Neonatal Cry Patterns in Infants with Cleft Lip and/or Palate Prior to Surgery: A Pilot Study“ angewendet werden (Kapitel 1.2.1). Daher beschreibt Kapitel 2 die zu analysierende Datenbasis des Projektes (Kapitel 2.1) sowie die Voranalyse der mir vorgegebenen Testsignale geeigneter Patientenaufnahmen aus dem Archiv des ZVES (Kapitel 2.3). Die Auswahl der anonymisierten Testsignale erfolgte nach bestimmten Kriterien: Es wurden Aufnahmen von Säuglingen und Kindern unterschiedlichen Alters aus unterschiedlichen Patientengruppen vorgegeben, in denen es typische akustische Phänomene gab, die die angestrebte semi-automatische Analyse vor besondere Herausforderungen stellte.

### 2.1 Projektdatenbasis und Messgrößen

Die hier analysierten pseudonymisierten Audiofiles stammten von den acht Probanden der Pilotstudie „Evaluation of the Neurobiology of Neonatal Cry Patterns in Infants with Cleft Lip and/or Palate Prior to Surgery: A Pilot Study“ (siehe Kapitel 1.2.1), für die das ZVES in Forschungskooperation die Säuglingslaute ausgewertet hat. Im Rahmen vorliegender Arbeit wurden temporale Lauteigenschaften dieses Datensatzes analysiert. Erste Ergebnisse dieser Analysen wurden auf dem Kongress der European Cleft Palate Craniofacial Association (ECPA) in Form eines Posters präsentiert (vgl. Eigene wissenschaftliche Tätigkeit) und in „Pediatric Research“ publiziert (Conrad et al., 2020).

In Tabelle 1 ist die Datenbasis der hier durchgeführten Projektanalysen an 274 Vokalisationen berichtet.

Tabelle 1: Projektdatenbasis für die Gruppe der Säuglinge mit (LKGS) und ohne (KG) oro-faziale Spaltbildungen

Proband	Gruppe	Alter in Tagen	Anzahl der auswertbaren Laute
1	LKGS	82	20
2	LKGS	53	12
3	LKGS	52	30
4	LKGS	62	14
5	KG	56	31
6	KG	58	67
7	KG	56	32
8	KG	69	68

Folgende temporale Messgrößen wurden mit Hilfe des zuvor erarbeiteten und getesteten PRAAT Skriptes (vgl. Kap. 1.2.1 und 1.2.2) ermittelt und auf Signallevel deskriptiv (im Methodenvergleich: SAM oder MM) berichtet:

**Gesamtdauer (Lges) der Einzellaute (SAM oder MM in [s])** (Vokalisationen innerhalb einer expiratorischen Phase) pro Melodiestrukturkategorie. Dabei wurden die typischen Strukturkategorien der vorsprachlichen Entwicklung analysiert: 1B, 2B, 3B, MB, 1S, KS, 2S, 3S, MS, ? (fuzzy class).

Diese Strukturkategorien beschreiben Säuglingslaute, deren Melodie nur aus einem Bogen besteht (1B) bzw. solche mit komplexer (mehrbögiger) melodischer Struktur (2B, 3B, MB), die zusätzlich durch innermelodische Pausen rhythmisch segmentiert sein können (alle \*S; für weitere Details siehe z.B. (Birr, 2009).

**Dauer der einzelnen Bögen (B [s])** in segmentierten Lauten (N= 1, 2, ...n).

**Dauer der einzelnen innerlautlichen Pausen (p [s])** in segmentierten Lauten (p12, p23, p34, p45...).

Die mithilfe des Skriptes ermittelten Werte schließen ebenfalls die **Inspirationslängen (I [S])** mit ein ( $N=1,2, \dots n$ ).

Alle Größen wurden pro Melodiestrukturkategorie auf Signallevel analysiert.

Auf Probandenebene wurden der Melodiekomplexitätsindex (**MCI**) (Birr, 2009; Wermke, Leising, & Stellzig-Eisenhauer, 2007) und der Segmentierungsindex (**SI**) (Wermke et al., 2011) berechnet und über die Gruppe gemittelt.

Die Ergebnisse der temporalen Analysen durch Skriptanwendung wurden automatisch in Exceltabellen gespeichert und nachfolgend mit Hilfe deskriptiver Statistiken berichtet. Der Gruppenvergleich wurde mit Hilfe nicht-parametrischer Tests vorgenommen, Verteilungseigenschaften in Form von Boxplotdiagrammen berichtet. Für die statistische Analyse wurde das SPSS-Paket (Version 25) verwendet.

Die Ergebnisse dieser Analysen wurden an den Principal investigator (A. Conrad) gesandt und in den Projektpublikationen (Conrad et al., 2020; Conrad, Wermke, Kuhlmann, & Eisenmann, 2019) sowie in der Folgeprojektantragsstellung verwendet.

## 2.2 Vergleich semi-automatische und manuelle Projektdatenanalyse

Die Projektdaten (Kapitel 2.1) sollten semi-automatisch quantitativ analysiert werden. Entsprechend der Zielsetzung vorliegender Arbeit (Kapitel 1.2.1) musste ein solches Analysesystem zunächst entwickelt, und vom Autor getestet und implementiert werden. Zusätzlich zu der finalen Datenanalyse für das Projekt wurden alle Daten auch mit der bisher am ZVES verwendeten manuellen Methode analysiert und die Ergebnisse beider Methoden miteinander verglichen (siehe Kapitel 3.3.2). Damit sollte sichergestellt werden, dass die Projektdaten mithilfe des hier entwickelten Skriptes auch korrekt und mit hoher Zuverlässigkeit ermittelt wurden. Dieser Vergleich stellt somit neben der Skripttestung an den Testsignalen einen weiteren Reliabilitätstest für die hier entwickelten Skripte dar. Die Forderung nach einer semi-automatischen, auf dem Zeitsignal basierenden, Analysemethode für das Forschungsprojekt ergab sich aus Vorläuferarbeiten, die

zum einen gezeigt hatten, dass Säuglinge mit LKGS eine häufiger segmentierte Melodiestructur aufweisen und zum anderen dadurch, dass an der Melodiestructur vermessene temporale Größen im Vergleich zur direkten Vermessung am Zeitsignal fehlerhaft sein können.

## 2.3 Beschreibung der Voranalyse der Testsignale

Die vorgegebenen Testsignale stammen von verschiedenen Patientengruppen: Säuglinge und Kinder mit LKGS, Sagittalnahtsynostosen und sensorineuraler Hörschädigung sowie aus „unauffälligen“ Kontrollgruppen. Die Testsignale enthalten Lautäußerungen aus unterschiedlichen vorsprachlichen Entwicklungsphasen: spontanes Weinen, Silbenbabbeln, Satz- und Wortproduktion. Auf die ermittelten akustischen Besonderheiten der Gruppen wird in den zugehörigen Unterkapiteln detailliert eingegangen. Die Testsignale werden dabei mit Hilfe von Schmalband-Frequenzspektren und zugehörigen Zeitbereichsdarstellungen charakterisiert. Um die Zuordnung der Signale im Ergebnisteil zu vereinfachen, sind den Testsignalen Buchstaben in alphabetischer Reihenfolge zugeordnet.

Bei allen in den Kapiteln 2.3.1 bis 2.3.4 dargestellten Abbildungen handelt es sich im oberen Drittel um Zeitbereichsdarstellungen und in den unteren zwei Dritteln um instationäre 4kHz- Schmalbandspektrogramme. Die Darstellungen wurden, wie am ZVES üblich, routinemäßig im Rahmen der Voranalyse vom Autor mit dem Computerized Speech Lab (CSL) 4500 erzeugt (KAY Pentax Medical).

Alle CSL-Zeitdarstellungen zeigen im oberen Drittel ein Oszillogramm des Lautes: auf der y-Achse die Amplitude, auf der x-Achse die Zeit in Sekunden. Die Zeitachse ist, je nach Länge des Testsignals, unterschiedlich skaliert, da das CSL-System keine manuelle Skalenänderung erlaubt (immer automatische Zeitsignaldauer).

In den unteren zwei Dritteln werden auf der x-Achse wieder die Zeitachse, auf der y-Achse die Frequenzkomponenten des Signals linear bis 4kHz gezeigt. Die unterste Linie stellt die Grundfrequenz  $f_0$  des Lautes dar, dessen zeitlicher Verlauf als der Melodieverlauf des Lautes bezeichnet wird. Die darüber liegenden

Frequenzbänder werden als Harmonische bezeichnet. Sie sind ganzzahlige Vielfache der Grundfrequenz  $f_0$ . Die Intensität der Frequenzanteile wird durch deren Schwärzungsgrad dargestellt, je intensiver diese geschwärzt sind, desto höher die Intensität des jeweiligen Frequenzbandes.

### 2.3.1 Testsignale aus der Patientengruppe der Säuglinge und Kinder mit LKGS-Spaltbildung

Als Testsignale aus der Patientengruppe der Kinder mit Lippen- Kiefer- Gaumen und Segelspalten wurden vier charakteristische Langserien vorgegeben.

Die ersten beiden Langserien stammen von einem weiblichen Säugling im Alter von 68 Tagen in der Kondition mit bzw. ohne eingesetzter Oberkieferplatte<sup>2</sup> (vgl. auch Hotz & Gnoinski, 1979; Komposch, 1986). Das Kind hatte eine nichtsyndromale, isolierte Gaumenspalte (HSH nach LAHSHAL-Klassifikation). Je eine Aufnahme wurde mit beziehungsweise ohne eingesetzte Oberkieferplatte angefertigt. Die Oberkieferplatte wurde in der interdisziplinären Spaltsprechstunde an der Poliklinik für Kieferorthopädie am UKW Würzburg (Direktorin: Prof. Dr. Angelika Stellzig-Eisenhauer) angefertigt und eingeschliffen. Die beiden Aufnahmen wurden im Rahmen dieser Sprechstunde zeitlich unmittelbar nacheinander durchgeführt.

---

<sup>2</sup> Im Rahmen der Kieferorthopädischen Frühtherapie bei Säuglingen mit LKGS werden Oberkieferplatten angefertigt. Sie dienen in erster Linie zur Ausformung und Harmonisierung des gespaltenen Kiefers (u.a. indem sie die Zunge aus dem Spalt fernhalten) vor einer korrekativen Operation.



### 2.3.1.1 Testbeispiel bei Gaumenspalte (HSH) ohne eingesetzte Oberkieferplatte

Bei **Testsignal A** handelt es sich um eine Sequenz von spontan erzeugten Weinlauten, die ohne eingesetzte Oberkieferplatte von einem Säugling im Alter von 68 Tagen erzeugt wurden (Abbildung 3).

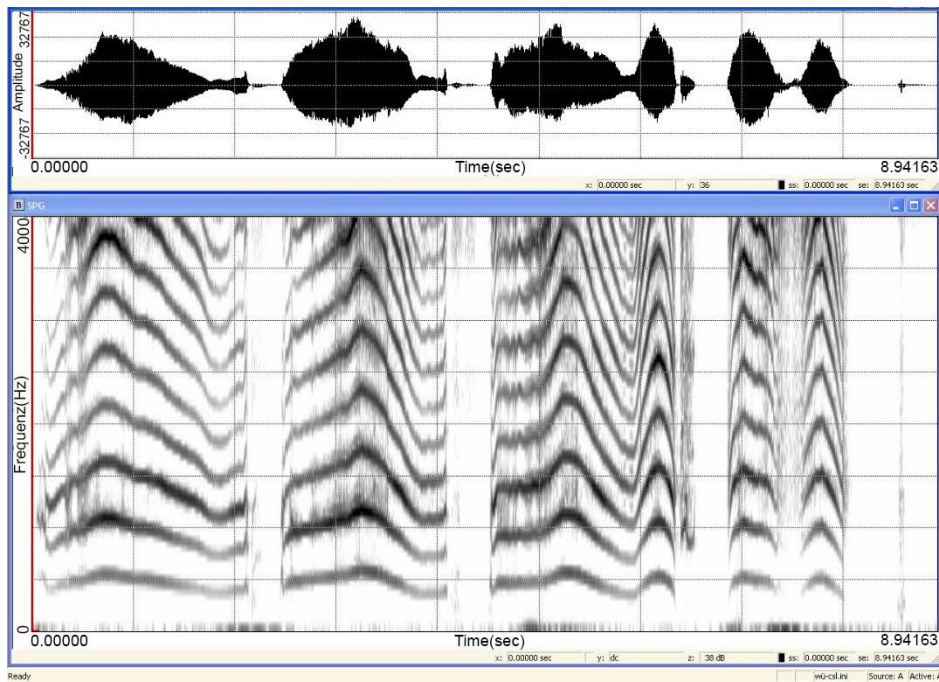


Abbildung 3: Testsignal A: Zeitbereichsdarstellung und Schmalband-Spektrogramm  
Sequenz mit Weinlauten [Mädchen, 68 Tage; isolierte Gaumenspalte ohne eingesetzte Oberkieferplatte]

Die Sequenz ist etwa 9 Sekunden lang und enthält vier phonierte Einzellaute (Weinlaute). Dazwischen finden sich „stille“ Bereiche, die auch die partiell geräuschartigen Inspirationen enthalten.

Die Melodie der beiden ersten Weinlaute ist einbögig, das heißt sie besteht im Wesentlichen aus einer ansteigenden  $f_0$  Flanke, gefolgt von einer absteigenden. Der erste Laut enthält zusätzlich sehr kurze Ein- und Ausschwingphasen. Die Melodiekonturen der ersten beiden Laute sind in ihrer Form unterschiedlich, wie man an den höheren Harmonischen sehr gut erkennt.

Die Melodie des dritten Lautes besteht aus einem Doppelbogen. Der vierte Laut enthält ebenfalls eine zweibögige Melodie, in der die beiden Bögen aber durch

eine sogenannte Segmentierungspause, hier mit deutlichem Geräuschanteil, getrennt sind. Es handelt sich um eine laryngeale Konstriktion (Rhythmisierung), wie sie insbesondere bei Säuglingen mit LKGS-Spaltbildungen häufig vorkommt (Ostermann, 2010).

### 2.3.1.2 Testsignal für eine Lautsequenz bei Gaumenspalte (HSH) mit eingesetzter Oberkieferplatte

Dieses **Testsignal B** vom selben Säugling (siehe Kapitel 2.3.1.1) enthält eine Sequenz spontaner Weinlaute aus der Aufnahme mit der, nach dem Einschleifen, wieder eingesetzten Oberkieferplatte (Abbildung 4).

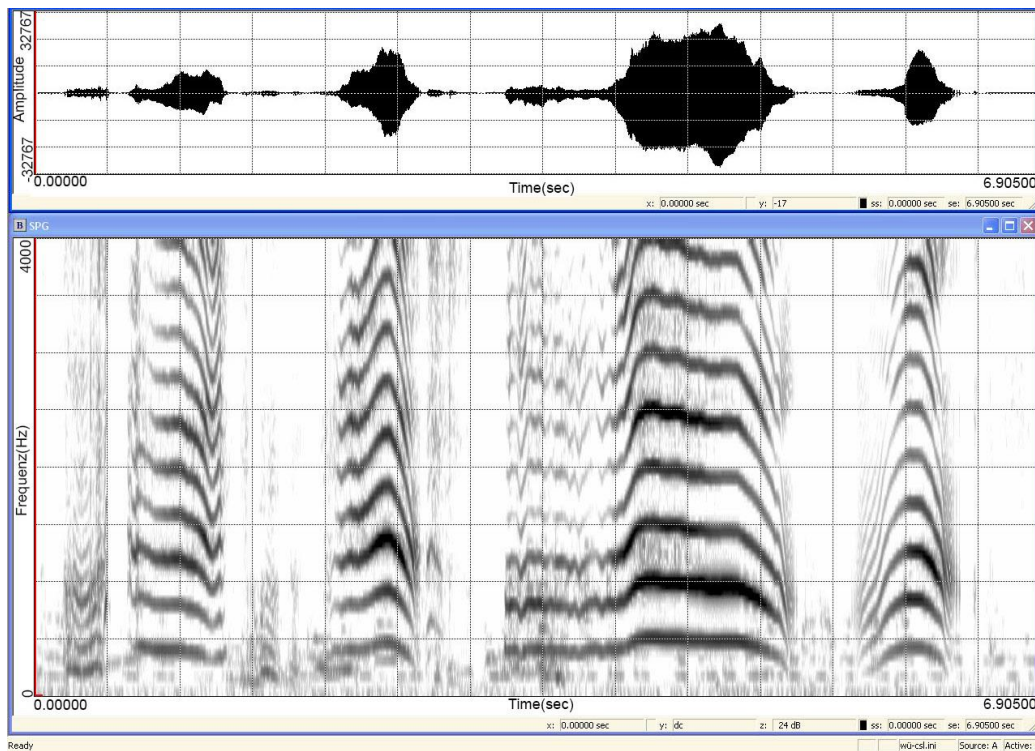


Abbildung 4: Testsignal B: Zeitbereichsdarstellung und Schmalband-Spektrogramm  
Sequenz mit Übergangslaute (mitigierte Weinlaute) [Mädchen, 68 Tage; isolierte Gaumenspalte HSH mit eingesetzter Oberkieferplatte]

Die Sequenz von 6,9 Sekunden Dauer enthält starke Intensitätsunterschiede der Laute, die man in der Zeitbereichsdarstellung erkennt; es gibt zwei sogenannte Übergangslaute (mitigierte, in ihrer Intensität abgeschwächte Weinlaute). Auch hier sind die Inspirationsgeräusche beziehungsweise inspiratorische Phonationen sehr deutlich zu erkennen. Die Sequenz enthält zusätzliche

Hintergrundgeräusche in Form unverständlicher Gespräche während der klinischen Behandlung. Die Aufnahme wurde während der Vorstellung des Säuglings in der interdisziplinären Spaltsprechstunde an der Poliklinik für Kieferorthopädie des UKW gemacht.

Die ersten beiden Laute sind mit geringerer Intensität als der dritte Laut phoniert. Während die ersten beiden Laute vergleichsweise kurz sind und die Melodie einmal fällt und im zweiten Fall ansteigt, ist der dritte Laut vergleichsweise lang und zeigt eine komplexe Melodie. Sie besteht aus einem tremoloartig variierenden Plateau, gefolgt von einem Melodiebogen, einer stimmlosen Segmentierung (glottal stop) und einem zweiten Bogen. Die Ränder der Melodie zur Segmentierung enthalten Konstriktionsgeräusche (Pressen). Hier gibt es eine starke Variabilität der Grundfrequenz und innerlautliche Intensitätsunterschiede: Der Weinlaut beginnt mit schnellen  $f_0$  Variationen bei ca. 403 Hz und steigt relativ schnell auf ca. 504 Hz an. Der bogenartige Verlauf zeigt eine steile finale Flanke, gefolgt von einer Segmentierungspause (420ms), der ein kurzer Bogen folgt. Während der Energiegehalt bei den ersten beiden Lauten auf der dritten Harmonischen am höchsten ist, ist im dritten Laut die zweite Harmonische besonders energiereich.

#### 2.3.1.3 Testsignale für Lautsequenzen bei doppelseitiger LKGS im Sprachalter

Die nachfolgenden beiden Langserien stammen von einem Jungen mit ehemals doppelseitiger LKGS-Spalte (LASHSAL-Klassifikation: LASHKaKL) im Alter von 800 respektive 926 Tagen, lange nach einer erfolgten kieferorthopädischen Frühbehandlung sowie operativem Lippen- und Gaumenverschluss in der Mund-Kiefer und Gesichtschirurgie der ZMK-Poliklinik in Würzburg (Direktor: Prof. Alexander Kübler). Das Kind befindet sich bereits im Sprachalter.

#### Testsignal für einen Dreiwortsatz

Beim **Testsignal C** (3,3s lang), das im Sprachalter (800 Tage) aufgenommen wurde, handelt es sich um einen Dreiwortsatz, bei dem das erste Wort wiederholt wird („Opa, Opa Bäume gespritzt“) (Abbildung 5). Der Satz stammt aus einem

Sprachtest am ZVES. Das Kind erzählt spontan von seiner Beobachtung der Gartenarbeit seines Großvaters. Auch in dieser Sequenz wechseln sich phonierte beziehungsweise artikulierte Bereiche mit Pausen ab; die Segmentierungspausen trennen hier sowohl Silben als auch Wörter. Die Aufnahmen erfolgten nicht in einer Schallkabine und enthalten daher Hintergrundgeräusche.

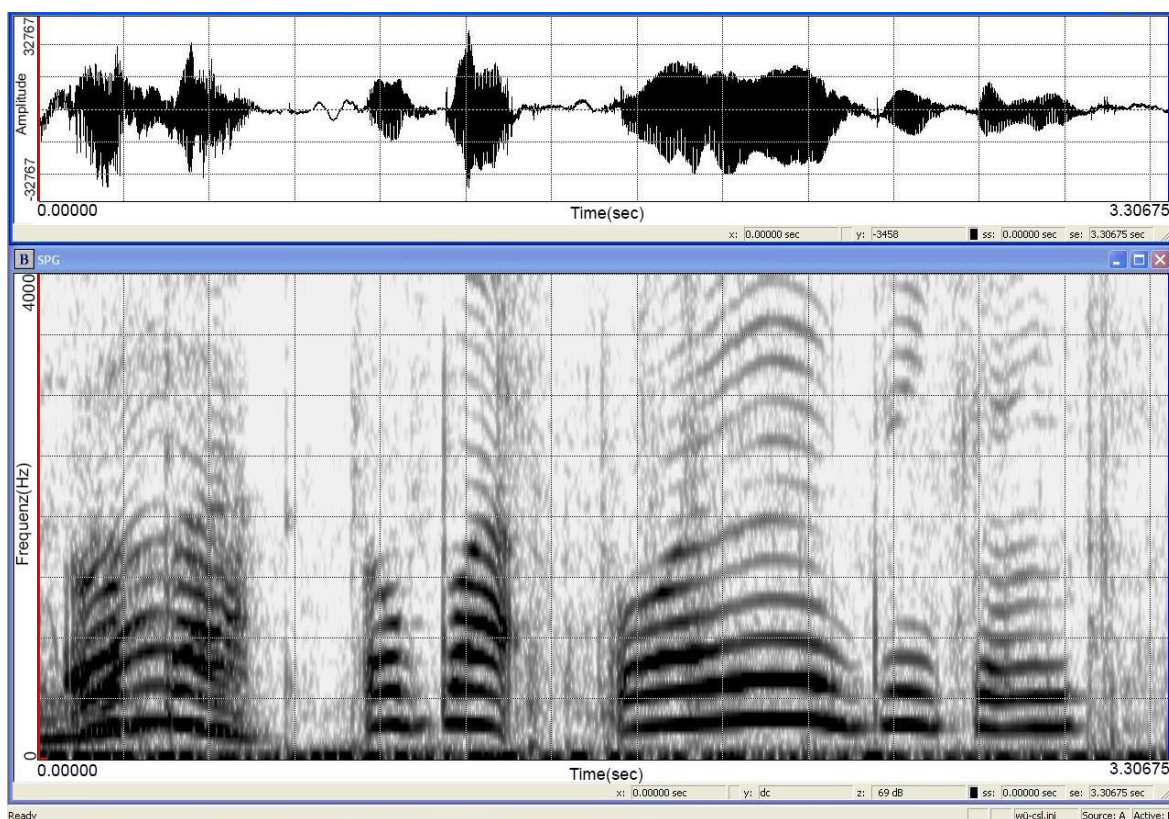


Abbildung 5: Testsignal C: Zeitbereichsdarstellung und Schmalband-Spektrogramm  
Dreiwortsatz („Opa, Opa Bäume gespritzt“) [Junge, 800 Tage; LASHKaKL]

Das Kind spricht teilweise gut verständliche Worte. Das Wort „gespritzt“ allerdings wird durch die Lautkombination „ge-si“ artikuliert, durch Konsonantenauslassung wäre es ohne Kontextinformationen aus dem Begleitprotokoll unverständlich. Signalanalytisch interessant an diesem Testsignal ist die Variation des Wortes „Opa“. Am Anfang der Sequenz wird es zunächst mit durchgängiger Melodiekontur (Einerbogen) artikuliert, während bei der Wiederholung des Wortes eine Pause zwischen den beiden Silben gelassen wurde („O-pa“ mit Segmentierungspause). Das Wort „Bäume“ zeigt einen

Melodiebogen mit steigender Kontur. Das Wort „gespritzt“ ist ebenfalls durch eine Segmentierungspause zwischen den Silben charakterisiert. Die Inspiration ist hier schwieriger zu identifizieren, da sie partiell im Hintergrundrauschen untergeht. Weiterhin imponieren, auch in diesem Beispiel, starke Intensitätsschwankungen innerhalb der Langserie.

### Testsignal für einen Vierwortsatz

Das **Testsignal D** zeigt einen Vierwortsatz des gleichen Kindes (Abbildung 6), vereinhalf Monate später: „Der Radler hat umfalle [ist umgefallen]“. Die Sequenz ist 2,55s lang. Im Vergleich zum Beispiel in Abbildung 5 ist hier die Sprechgeschwindigkeit erhöht. Durch noch vorhandene Artikulationsschwierigkeiten wird die Verständlichkeit des Satzes stark eingeschränkt.

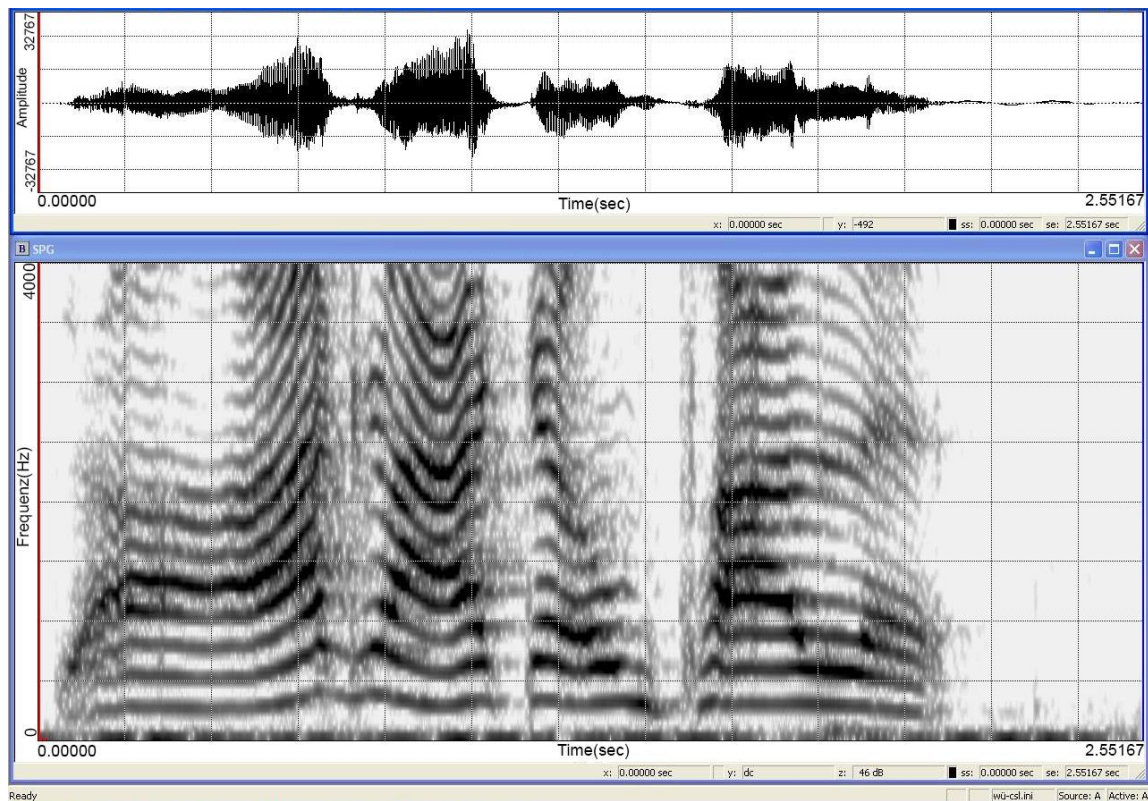


Abbildung 6: Testsignal D: Zeitbereichsdarstellung und Schmalband-Spektrogramm  
 Vierwortsatz („Der Radler hat umfalle [ist umgefallen]“) [Junge, 926 Tage; LASHKaKL]

#### 2.3.1.4 Testsignale aus der Patientengruppe der Säuglinge mit Sagittalnahtsynostosen

Die anonymisierten Testsignale stammen ursprünglich aus einem Forschungsprojekt (CFCW Projekt Ethikvotumnummer: 143/09), das interdisziplinär zwischen der Kieferorthopädie, Kieferchirurgie und pädiatrischen Neurochirurgie durchgeführt wurde. Inhalt dieses Pilotprojektes war es, Säuglinge mit Sagittalnahtsynostosen bezüglich ihres Kopfwachstums sowie ihrer kognitiven und motorischen Entwicklung zu untersuchen. Im Rahmen der Untersuchung früher kognitiver Fähigkeiten wurden auch Lautaufnahmen (Komfortvokalisationen in Mutter-Kind-Interaktionen: Babbeln) dieser Säuglinge zu festgelegten Altersbereichen angefertigt und im ZVES ausgewertet; im Fall der Säuglinge, die sich einer neurochirurgischen Operation unterziehen mussten, wurden Lautaufnahmen (Babbeln) prä- und postoperativ angefertigt (Linz et al., 2018; Vink; Wermke et al., 2017).

Die hier verwendeten Testsignale stammen jeweils von einem Mädchen und einem Jungen mit Sagittalnahtsynostose, sowie einem Probanden aus der Kontrollgruppe des Projekts.

##### Testsignal einer Lautsequenz bei Sagittalnahtsynostose

Beim nachfolgenden **Testsignal E** handelt es sich um eine Lautsequenz mit artikulierte(n) Komfortvokalisationen (Babbeln) eines Jungen mit Sagittalnahtsynostose (Abbildung 7).

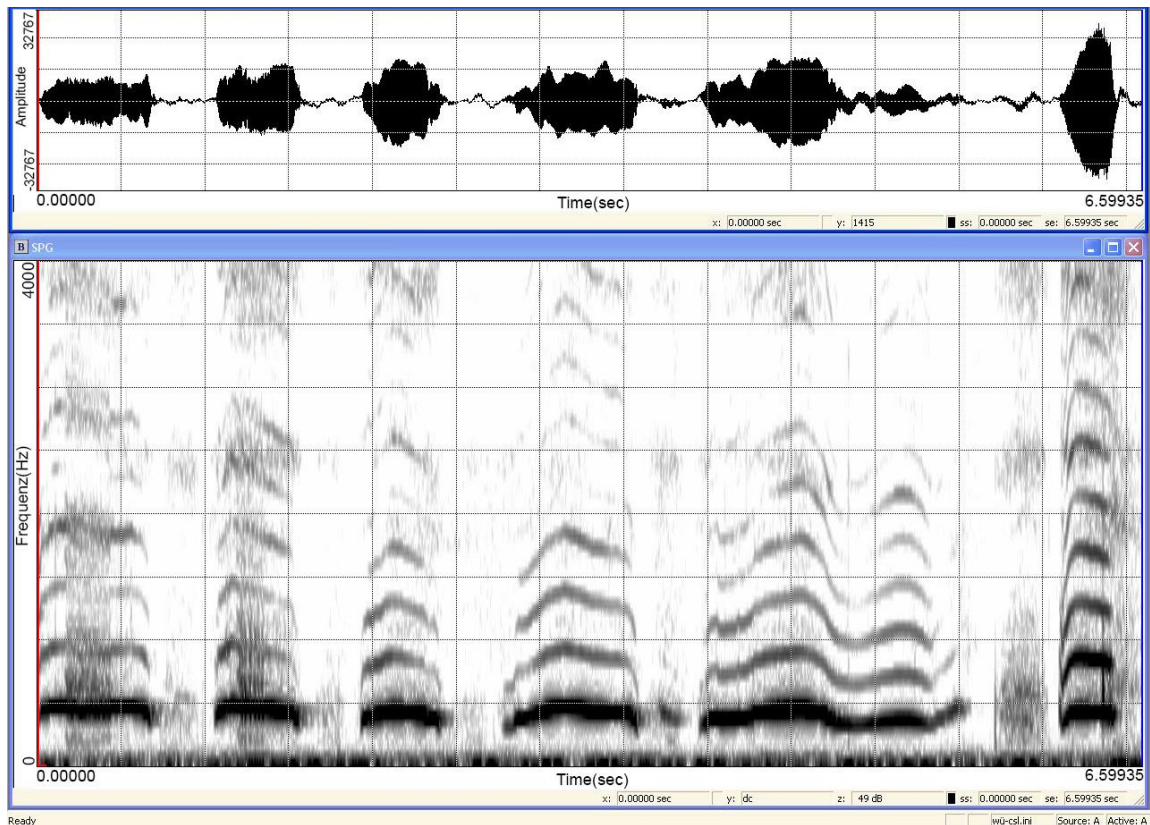


Abbildung 7: Testsignal E: Zeitbereichsdarstellung und Schmalbandspektrum Sequenz mit artikulierten Komfortvokalisationen (Babbeln) [Junge, 222 Tage; Sagittalnahtsynostose]

Die Sequenz (Abbildung 7) besteht aus sechs Einzellaute unterschiedlicher Länge (Gesamtlänge 6,59s). Dazwischen befinden sich Pausen, in denen die Inspiration erfolgt (teilweise hörbare Inspirationslaute, die im Spektrum als Rauschbänder erkennbar sind). Die ersten beiden Einzellaute enthalten im Mittelteil eine geräuschartige Komponente (phonatorisches Rauschen). Der fünfte Laut besteht, anders als die vier vorausgehenden einböigen Laute, aus einer komplexen doppelböigen Melodie, gefolgt von einem finalen, ausgeprägten fo-Gliding (kurze, steigende Flanke am Lautende; siehe Abbildung 7). Der Intensitätsgehalt der Grundfrequenz ist im Vergleich zu allen übrigen Harmonischen der Höchste. Ab ca. 2000 Hz sind die Frequenzkomponenten deutlich intensitätsärmer.

## Testsignal einer Lautsequenz bei Sagittalnahtsynostose (präoperativ)

Auch bei diesem **Testsignal F** handelt es sich um eine Abfolge von artikulierten Komfortvokalisationen (Babbeln), hier produziert von einem Mädchen mit Sagittalnahtsynostose im Alter von 255 Tagen (Abbildung 8). Die Aufzeichnung erfolgte präoperativ und hat eine Gesamtlänge von 4,45s.

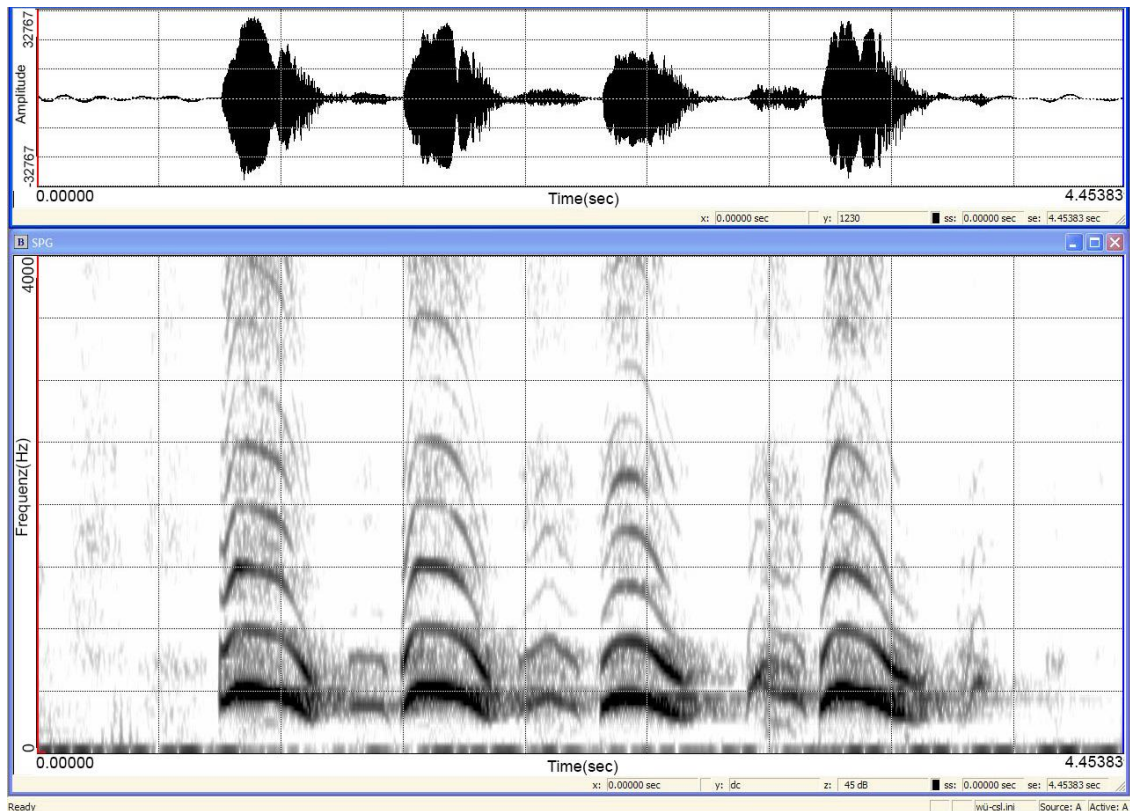


Abbildung 8: Testsignal F: Zeitbereichsdarstellung und Schmalbandspektrogramm Sequenz mit artikulierten Komfortvokalisationen (Babbeln) [Mädchen, 255 Tage; prä OP; Sagittalnahtsynostose]

Im Unterschied zu dem in Abbildung 7 beschriebenen Testsignal enthält das in Abbildung 8 dargestellte Signal nur einzelbögige, einfache Babbellaute (vier Stück). Diese sind hier durch sehr markante, melodisch ebenfalls bogenförmige Inspirationsphonationen unterbrochen. Der Intensitätsgehalt ist auch hier im Bereich der Grundfrequenz am höchsten; aber auch die Frequenzbereiche der zweiten und dritten Harmonischen verfügen noch über hohe Energiedichten (Schwärzung).



### 2.3.2 Testsignal für Babbellaute eines Säuglings aus der Kontrollgruppe

Auch dieses 4,95s lange **Testsignal G** enthält eine Sequenz von artikulierte Komfortvokalisationen (Babbeln). Es stammt von einem Mädchen im Alter von 211 Tagen aus der Kontrollgruppe des genannten CFCW-Forschungsprojektes.

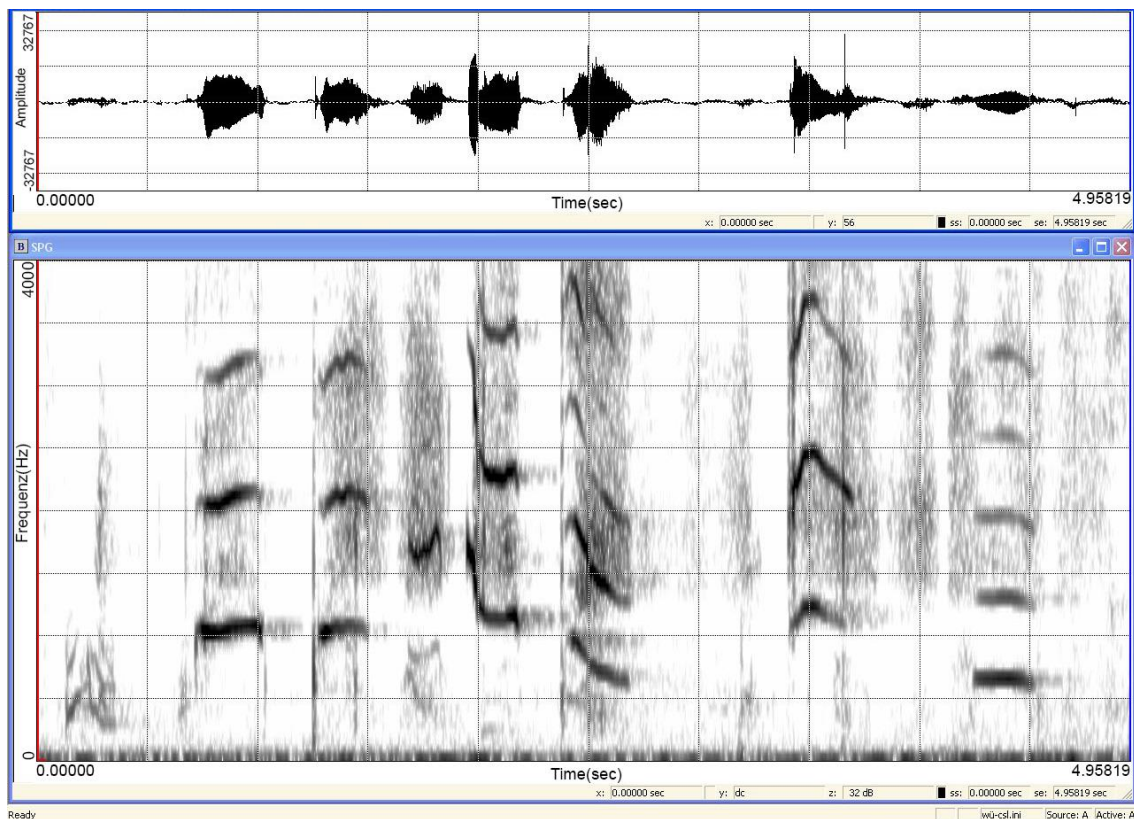


Abbildung 9: Testsignal G: Zeitbereichsdarstellung und Schmalbandspektrogramm Sequenz mit artikulierte Komfortvokalisationen (Babbeln) [Mädchen, 211 Tage; Kontrollgruppe „unauffällig“]

Dieses Testsignal (Abbildung 9) ist ausgewählt worden, da die Grundfrequenz der geäußerten Babbellaute deutlich erhöht ist. Hier handelt es sich um relativ hochfrequente Laute, sogenannte „Squeals“ (Buder, Warlaumont, & Oller, 2013). Diese Laute sind typische Vokalisationen für bestimmte altersabhängige Babbelphasen unauffälliger Säuglinge (Pachtner, 2017). Zusätzlich enthält die Sequenz teilweise sehr schnelles Gliding (Glissando) über 1000 Hz. Darüber hinaus imponieren hier die insgesamt auffällig kurzen Expirationslaute und ein deutlich segmentierter Expirationslaut am Anfang der abgebildeten Sequenz (Abbildung 9). Die Grundfrequenz variiert sehr stark und zeigt einen fo-Bereich

zwischen 600 und 2000 Hz. Das Mädchen phoniert vorwiegend im „Kopfreger“ (falsetto) und „trainiert“ zusätzlich die Kontrolle der Frequenzvariation beim Wechsel zwischen verschiedenen Produktionsregistern.

### 2.3.3 Testsignale aus der Patientengruppe der Säuglinge mit sensorineuraler Hörbeeinträchtigung oder -störung

Die beiden Testsignale aus den Aufnahmen der Patientengruppe der Säuglinge mit sensorineuralen Hörstörungen stammen von einem monozygoten Zwillingenpaar (Shehata-Dieler & Wermke, 2017). Eine der beiden Aufnahmen wurde im Alter von 79 Tagen vor Versorgung mit Hörhilfen durchgeführt. Die zweite Aufnahme wurde nach Versorgung mit Hörhilfen beim Geschwisterkind im Alter von 171 Tagen durchgeführt.

#### Testsignal einer Sequenz von Weinlauten vor Versorgung mit Hörhilfen

Das **Testsignal H** (Gesamtlänge 13,13s) enthält eine Sequenz mit Weinlauten eines Patienten des Comprehensive Hearing Center (CHC) der Klinik und Poliklinik für Hals-, Nasen- und Ohrenkrankheiten, plastische und ästhetische Operationen in Würzburg (Direktor: Prof. Rudolf Hagen) im Alter von 79 Tagen vor der Versorgung mit Hörhilfen (Abbildung 10). Die Hörschwelle lag zu diesem Zeitpunkt bei 100dB HL (Shehata-Dieler & Wermke, 2017).

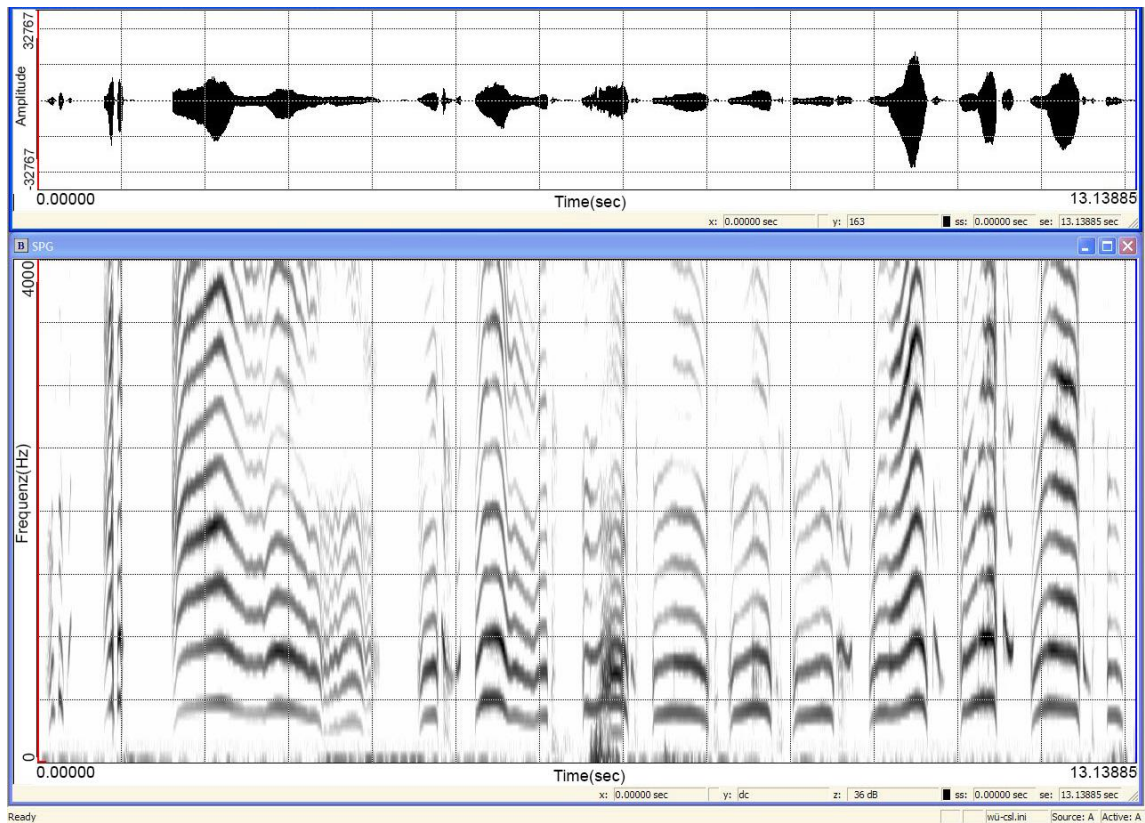


Abbildung 10: Testsignal H: Zeitbereichsdarstellung und Schmalband-Spektrum  
Sequenz mit Weinlauten [79 Tage, vor Versorgung mit Hörhilfen]

Typisch für sensorineural hörgeschädigte Säuglinge ist, unter anderem, die in der Zeitbereichsdarstellung in Abbildung 10 erkennbare Lautierung mit sehr starken Intensitätsschwankungen (Lautstärkeschwankungen). Die Melodie der Einzellaute zeigt verschiedene Mustervariationen (Einer- Zweier- und Dreierbögen in verschiedenen Konturformen und Kurzlaute, deren Länge <300ms sind). Im ersten längeren, dreibögigen Laut finden sich Anzeichen für phonatorische Regelstörungen, die sich in Regelzacken und tremoloartigen Phänomenen zeigen. Die egressiven Laute variieren in der Dauer und die Inspirationen sind akustisch in markante, rhythmische Phrasen eingebaut. Im fünften Laut erkennt man Subharmonische.

Testsignal einer Sequenz von Komfortlauten nach Versorgung mit Hörhilfen

**Testsignal I** (Abbildung 11) enthält Komfortlaute, in diesem Fall sehr frühe Babbellaute (Cooing). Es handelt sich um das Geschwisterkind des Säuglings von Testsignal H (siehe Kapitel 2.3.3 und Abbildung 10). Der Säugling trug bei der Aufnahme beidseits eine Hörhilfe: Die in Abbildung 11 beschriebene Lautserie ist mit aktivierten Hörgeräten geäußert worden (Gesamtlänge 9,2s).

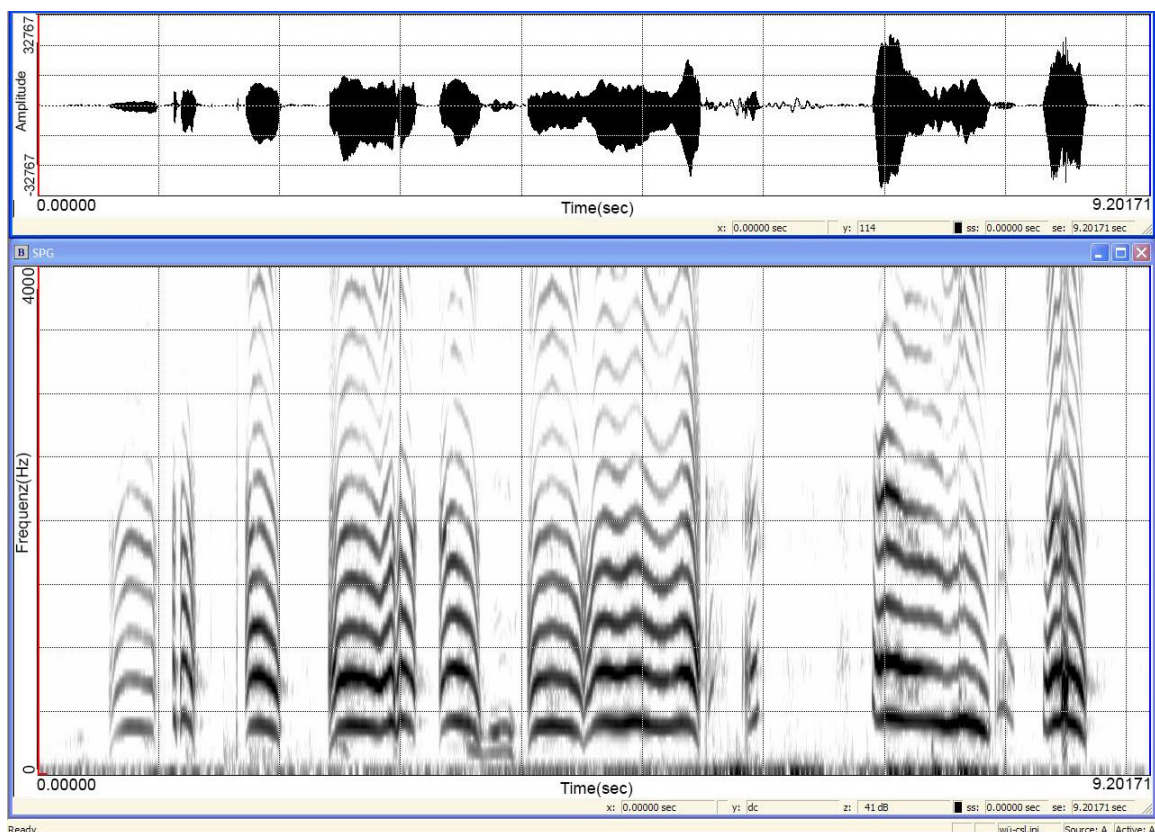


Abbildung 11: Testsignal I: Zeitbereichsdarstellung und Schmalbandspektrogramm Sequenz mit Komfortvokalisationen: sehr frühe Babbellaute (Cooing) [171 Tage; mit Hörhilfen]

Auch in dieser Aufnahme variiert die Melodie und es kommen innerlautliche Segmentierungspausen vor. Besonders im ersten Laut der Sequenz ist die innerlautliche Segmentierung deutlich ausgeprägt. Die innerlautliche Segmentierungspause im ersten Laut hat fast die gleiche Länge wie die darauffolgende Zwischenlautpause (vgl. mit Abbildung 42).

### 2.3.4 Testsignale aus der Entwicklungsphase des Kanonischen Babbels

Das kanonische Babbeln ist die vorsprachliche Entwicklungsphase, die der Wortproduktion unmittelbar vorausgeht. Sie ist charakterisiert durch Silbenketten, die aus wohlgeformten Konsonant-Vokalverbindungen bestehen (Buder et al., 2013). Die akustischen Eigenschaften der Silben, insbesondere die Vokalformanten und die Formantübergangszeiten in Konsonant-Vokal-Elementen, entsprechen bereits weitgehend denen der Zielsprache (hier Deutsch). Die Silben innerhalb einer Äußerung können sowohl vollständig (Abbildung 12) als auch unvollständig segmentiert sein.

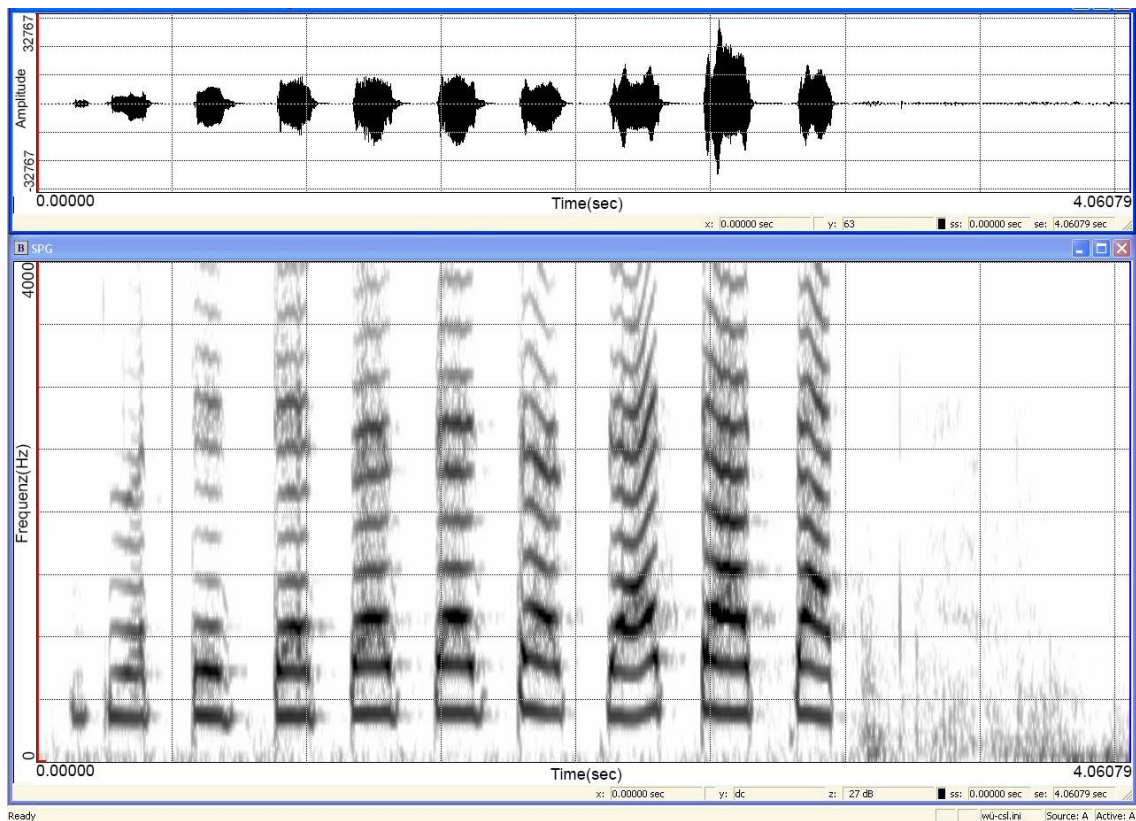


Abbildung 12: Zeitbereichsdarstellung und Schmalbandspektrum einer Sequenz kanonischer Babbelsilben, die durch vollständige Segmentierungen getrennt sind

**Testsignal J** (Abbildung 13; 4,06s) zeigt eine unvollständige Segmentierung innerhalb der mehrsilbigen Einzelsignale [ma-ma-ma]. Aufgrund der nasalen Konsonanten [m] ist die Grundfrequenz der Einzellaute hier durchgehend vorhanden.

Obwohl es sich um dieselbe Silbenfolge wie in Abbildung 12 [ma-ma] handelt, wird die Melodiekontur der vier Einzellaute in Abbildung 13 (Gesamtlänge 5,49s) variiert. Die Laute sind durch Pausen, in denen die Inspiration erfolgt (Rauschbänder), getrennt.

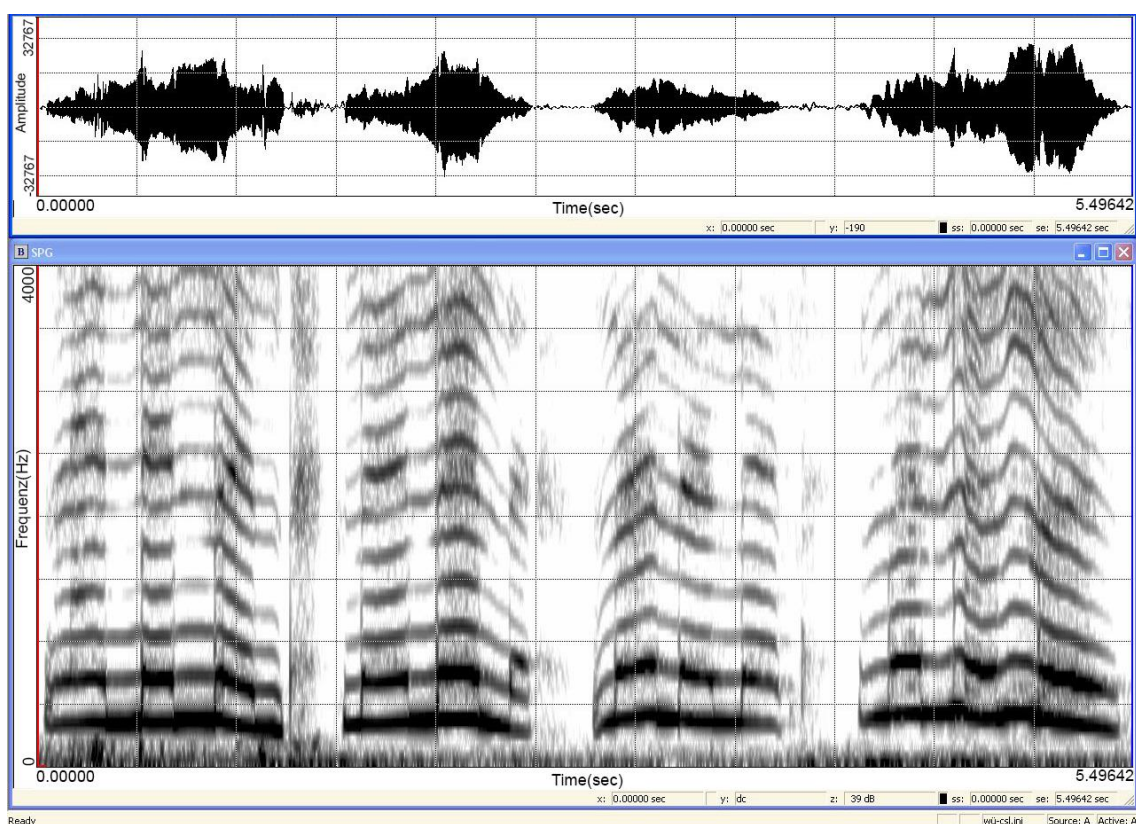


Abbildung 13: Testsignal J: Zeitbereichsdarstellung und Schmalband-Spektrogramm Sequenz von kanonischen Babbellauten mit unvollständiger Segmentierung innerhalb der Einzelsignale [Kind, 280 Tage]

Es gibt Intensitätsschwankungen innerhalb der Einzellaute zwischen Vokalen [a] und Konsonanten [m] bei durchgehender  $f_0$ -Kontur.

## 2.4 Kurzbeschreibung der Analysesoftware PRAAT

### 2.4.1 Grafische Oberfläche PRAAT

Die grafische Oberfläche besteht aus drei Hauptfenstern (PRAAT Objects, Picture und Info), über die sich die einzelnen Funktionen und Befehle aufrufen, verarbeiten beziehungsweise darstellen lassen. Sie sollen im Folgenden kurz vorgestellt werden. Die nachfolgenden Abbildungen enthalten exemplarisch Darstellungen einer Windowsversion (Microsoft Windows 7 x64) von PRAAT. Die Oberflächen auf anderen Betriebssystemen weichen leicht ab.

Vom PRAAT Objects Window (Abbildung 14) lassen sich u.a. Objekte wie Sounddateien, Textdateien (die z.B. Annotationen ausgewählter zeitlicher Abschnitte enthalten) und von PRAAT selbst erzeugte Objekte laden, speichern, exportieren und manipulieren. Je nach geladenem und durch Anklicken aktiv geschaltetem Objekt, stehen verschiedene Optionen zur Verarbeitung des eingeladenen Objekts zur Verfügung; werden mehrere Objekte gleichzeitig angewählt, stehen dem Nutzer, je nach Kombination, weitere Optionen zur Weiterverarbeitung zur Verfügung. Abbildung 14 zeigt diese Optionen exemplarisch für eine eingeladene und dann angeklickte Sounddatei im Wave-Format, die eine Lautfolge aus den vorgegebenen Testsignalen enthält. Die blau hinterlegte Datei im Objektfenster ist durch Anklicken aktiv geschaltet worden. Auf der rechten Seite befinden sich die nach Aktivierung für diese Dateiverfügbaren Optionen. Die für die vorliegende Arbeit relevanten Optionen werden nachfolgend im Zusammenhang mit ihrer Anwendungsfunktion beschrieben.

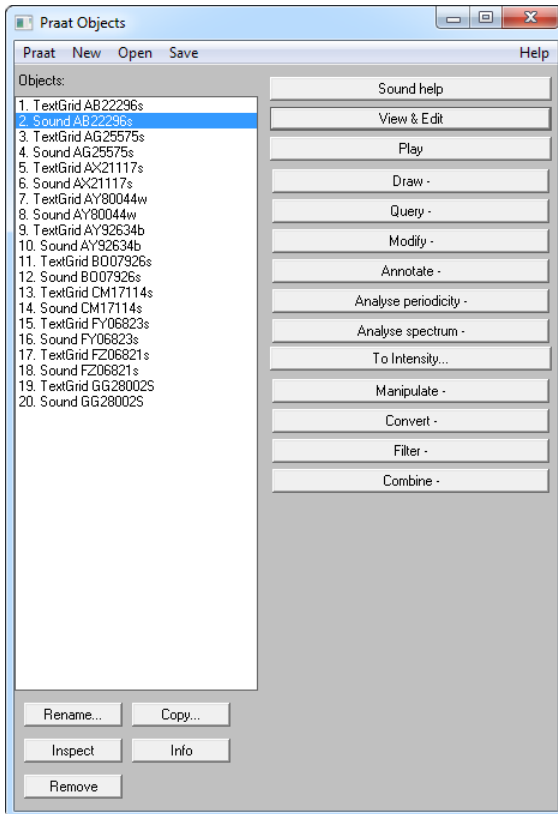


Abbildung 14: Praat Objects Fenster mit eingeladenen Sound- und Textgriddateien  
Die zweite grafische Hauptoberfläche stellt das Praat Picture Window dar (Abbildung 15, links).

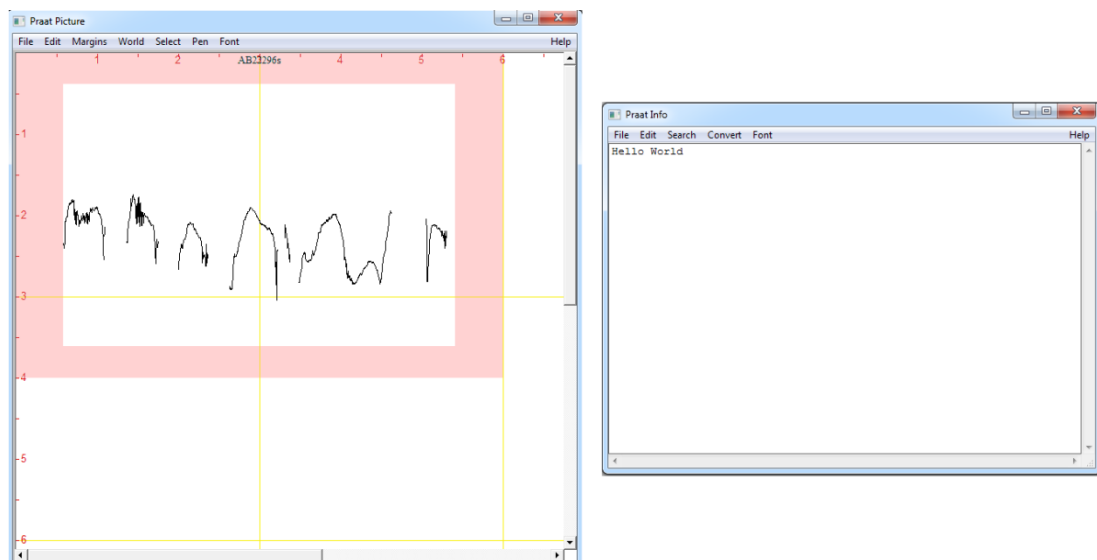


Abbildung 15: links: Praat Picture Window mit Viewport: Fenster mit Skalierung zum Ausgeben von grafischen Darstellungen: hier mit Pitchkonturen; rechts: PRAAT Info Window: einfaches Textfenster zum Ausgeben von Werten oder Texten



Das PRAAT Picture Window (Abbildung 15, links) dient zur grafischen Darstellung von in PRAAT erzeugten oder eingeladenen Objekten. Es lassen sich beispielsweise Grundfrequenz- oder Intensitätskurven grafisch darstellen. Hier ist exemplarisch die Grundfrequenzkurve der in Abbildung 14 aktivierten Sounddatei dargestellt. Am oberen Fensterrand stehen in Dropdownmenüs weitere Einstellungen zur grafischen Ausgabe zur Verfügung. Der rosafarbene Bereich („Viewport“ Abbildung 15) legt fest, in welchem Bereich die gewünschte Grafik ausgegeben werden soll.

In das PRAAT Info Window (Abbildung 15, rechts) können Informationen und Daten in Textform ausgegeben werden, die entweder von der grafischen Oberfläche oder von einem Skript aus angefordert werden. Auch hier stehen am oberen Rand des Fensters in Dropdownmenüs noch weitere Optionen zur Verfügung.

#### 2.4.1.1 Editorumgebung

Die Editorumgebung dient unter anderem dazu Sounddateien, Annotationen und Spektrogramme anzuzeigen, zu bearbeiten, zu annotieren oder an das PRAAT Picture Window (Abbildung 15) zur Erstellung von Grafiken und Kurven weiterzugeben. Sie kann, wenn eine Sounddatei eingeladen ist, vom PRAAT Objects Window aufgerufen werden.

Die Editorumgebung (Abbildung 16) besitzt eine gesonderte interne Programmumgebung, die von der Shellumgebung der PRAAT Picture und Objects Windows abzugrenzen ist.

Es wird pro eingeladener Sounddatei eine neue, nur der entsprechenden Sounddatei zugehörige, interne Programmumgebung erstellt. Falls in einem Skript Befehle für die jeweilige Editorumgebung einer Sounddatei ausgeführt werden sollen, muss per Befehl „*editor*:“ in diese gewechselt werden.

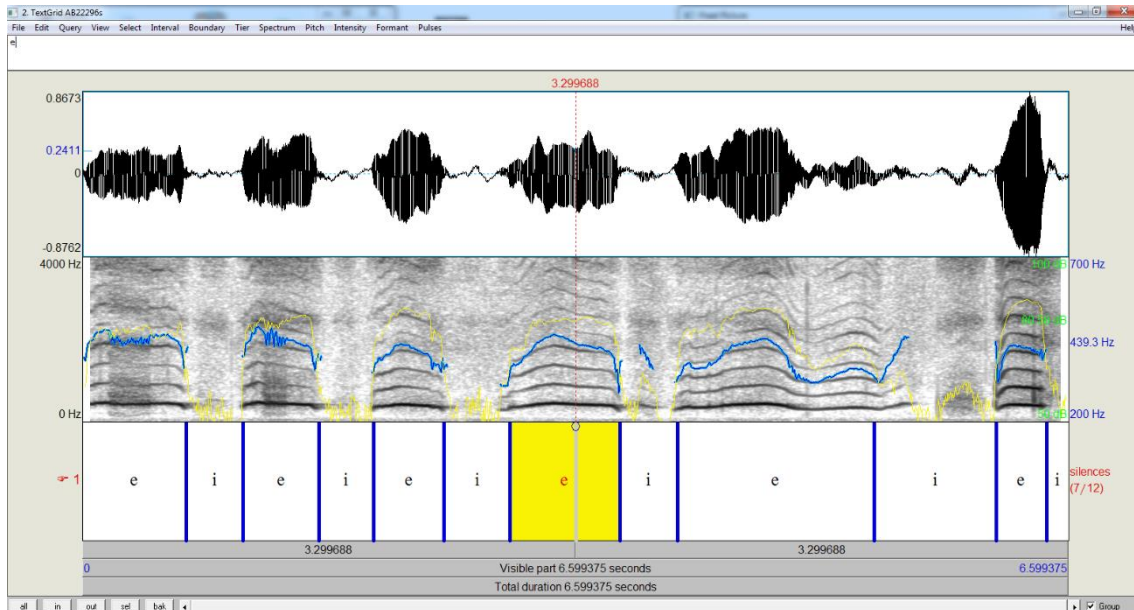


Abbildung 16: PRAAT Editorumgebung  
 Die Abbildung zeigt die Editorumgebung mit Zeitdarstellung, Frequenzspektrogramm mit eingezeichneter Intensitätskurve (gelb) und fo (blau) sowie eine Annotation in Lauten (e) und lautlose Anteile (i)

Auf den in Kapitel 3.2 (Ergebnisse der Skripttestung) beschriebenen Abbildungen (Abbildung 24 bis Abbildung 44) ist die PRAAT-Editorumgebung zu sehen. Im oberen Drittel befindet sich jeweils die Zeitbereichsdarstellung. In der Mitte befindet sich immer ein Schmalbandspektrogramm mit Frequenzbereich von 0-4000 Hz (lineare Skala). Die blaue Linie innerhalb des Spektrogramms zeigt den Verlauf der Grundfrequenzkurve (fo-Frequenzskala rechts), die gelbe Kurve den Verlauf der Intensität in dB (Skala rechts). Unten befindet sich das Annotationstier: die mit „i“ (Inspiration) beschrifteten Intervalle enthalten die „lautlosen“ oder inspiratorischen Anteile, die mit „e“ (Expiration) beschrifteten Intervalle die Anteile mit hier interessierenden phonierten Lauten. Die im Annotationstier gelb hinterlegten Elemente haben keine gesonderte Bedeutung: Es wird immer der Bereich hinterlegt, in dem zuletzt mit dem Mauszeiger geklickt wurde bzw. standardmäßig zunächst in der Signalmitte.

## 2.4.2 Textgrids

PRAAT stellt dem Nutzer über sogenannte TextGrid-Objekte die Möglichkeit zur Verfügung, ausgewählte Abschnitte einer Sounddatei zu beschriften. Dadurch hat man die Möglichkeit, eine oder mehrere unterschiedliche Aufteilungen zeitlicher Abschnitte vorzunehmen. Es können mehrere verschiedene zu beschriftende Abschnitte (Intervalle) für dieselbe Sounddatei angelegt werden, welche durch Grenzen („Boundaries“) voneinander getrennt sind. Diese können auch auf verschiedenen Ebenen (Tiers) angelegt werden.

Nach dem Einladen eines Soundobjektes in das PRAAT-Objects-Window, kann der Nutzer durch das Klicken auf die Schaltfläche „View & Edit“ (vgl. Abbildung 14) in die Editorumgebung gelangen, um dort die Annotation der jeweiligen Laute oder anderer inhaltlich und zeitlich relevanter Abschnitte unter Zuhilfenahme der Zeitbereichsdarstellung und eines Frequenzspektrogrammes vorzunehmen.

## 2.4.3 Die PRAAT Skriptsprache

PRAAT (Boersma & Weenink, 2018a) stellt eine eigene Skriptsprache zur Verfügung, die es dem Benutzer ermöglichen soll, häufig ausgeführte oder wiederkehrende Programmabläufe automatisiert durchzuführen. Alle Befehle, die man über PRAAT's grafische Oberfläche erreicht, können auch in Form eines Befehls in ein Skript eingebaut werden. Es können dabei nicht nur makroartige Skripte, die einzelne Befehle strikt nacheinander abarbeiten, erstellt werden: PRAAT's Skriptsprache ermöglicht darüber hinaus durchaus komplexere Programmroutinen durch den Einbau von Schleifen, Conditionals, Eingabeformen, Zuweisen von Variablen und weiteren auch in anderen Programmiersprachen geläufigen Programmkonstruktionen.

Der Programmierung des Skriptes ging die Einarbeitung in die programmeigene Skriptsprache voraus. Neben der offiziellen Bedienungsanleitung der Entwickler (Boersma & Weenink, 2018c) wurden hier auch die von einigen Nutzern selbst erstellten Anleitungen und Scripting-Tutorials verwendet. Besonders hilfreich waren hier ein inoffizielles Handbuch und ein Scripting Tutorial von Mayer (2017, 2018), auf deren Programmierbeispiele bei der Skripterstellung teilweise

zurückgegriffen wurde. Auch Onlineuser-Gruppen sowie Audit von Skripten anderer Benutzer halfen beim Einarbeiten in die Sprache und der Erstellung der Skripte. Der Autor verweist auf diese Quellen für detailliertere Beschreibungen, da hier nur solche methodischen Aspekte kurz erwähnt sind, die für die vorliegende Arbeit relevant waren.

## Wichtige Konstrukte der Skriptsprache

Es folgt eine Einleitung zu wiederkehrenden Konstrukten der PRAAT Skriptsprache. Diese finden in den im Ergebnisteil beschriebenen Skripten besonders häufig Anwendung und sollen das Verständnis der Skripte erleichtern.

### String- und andere Variablen

Zahlen sowie Kombinationen aus Zeichen lassen sich zur späteren Verwendung im Skript als Variablen speichern. Hierbei unterscheidet man zwischen Variablen, die Zahlenwerte speichern und Stringvariablen, bei denen beliebige Zeichenfolgen gespeichert werden können.

### Eingabeformen

Der Beginn einer Eingabeform wird durch das vorbelegte Wort „*form*“ gekennzeichnet, ihr Ende durch „*endform*“. Mit diesem Konstrukt lassen sich Eingabefenster ausgeben, die Eingaben und Einstellungen durch einen Benutzer erlauben und bereits Variablen zur späteren Verwendung speichern.

### Loops (Schleifen)

Soll ein bestimmter Codeabschnitt mehrfach wiederholt werden, wird dieser durch das vorbelegte Wort „*for*“ gekennzeichnet, gefolgt von einem initialen Zählerwert (z.B. „1“) und der Maximalanzahl der Wiederholungen für diesen Abschnitt. Die maximale Anzahl an Wiederholungen kann auch zuvor in einer Variablen abgelegt worden sein. Das Ende eines Loops wird schließlich mit „*endfor*“ gekennzeichnet. Falls die Wiederholung bis zum Eintreten eines bestimmten Ereignisses erfolgen soll, wird sie noch um „*until*“ sowie eine davon abhängige Bedingung ergänzt.

## Conditionals

Sollen Codeblöcke nur dann ausgeführt werden, wenn bestimmte Bedingungen eintreten bzw. erfüllt sind, bedient man sich der Conditionals. Der Beginn wird mit dem vorbelegten Wort „*if*“ gekennzeichnet, gefolgt von der Bedingung, z.B. ein bestimmter Wert für eine Variable. Sollen weitere oder andere Bedingungen angegeben werden, so kennzeichnet man diese mit „*elsif*“, gefolgt von der Bedingung. Beendet wird der Codeblock schließlich mit „*endif*“. Die Codeblöcke werden nur ausgeführt, falls die Bedingung, die hinter „*if*“ respektive „*elsif*“ folgt, erfüllt wurde.

## 3 Ergebnisse

Der Ergebnisteil besteht aus drei Teilbereichen. In **Teil 1** (Kapitel 3.1) werden die im Rahmen der Arbeit auf Basis entwicklungsdiagnostisch und medizinisch relevanter Aspekte vom Autor erarbeiteten Analyseroutinen (PRAAT-Skripte) ausführlich dargestellt. In **Teil 2** (Kapitel 3.2) wird die Testung der ProgrammROUTINEN an vorgegebenen, anonymisierten Testsignalen von Patientenaufnahmen im klinischen Umfeld dargestellt. Zusätzlich wird das erstellte Zusatzmodul „Timetable“ erläutert und systemvergleichend temporale Daten berichtet, die für die Vorsprachliche Diagnostik bei Säuglingen mit LKGS-Spaltbildung besonders relevant sind. In **Teil 3** (Kapitel 3.3) sind die Ergebnisse der Anwendung des Skriptes auf die Projektdaten berichtet.

### 3.1 Teil 1: Aufbau und Testung des erarbeiteten Skriptes

Der finale Aufbau des hier erarbeiteten Skriptes ist in Abbildung 17 übersichtlich dargestellt, die einzelnen Codeabschnitte des Skriptes werden in den nachfolgenden Kapiteln erklärt und durch detaillierte Informationen im Anhang A 1 ergänzt.

In einer grafischen Eingabeform nimmt der Benutzer Voreinstellungen vor. Nach dem Erstellen eines Ausgabeordners sowie einer Dateiliste durch das Skript, werden Sounddateien und zugehörige Textgrids (Annotation der Sounddateien mithilfe des Skriptes in Kapitel 3.1.2) sukzessive geladen, ausgewählte Passagen extrahiert und den gewünschten Analysen zugeführt. Die Ergebnisse werden in Tabellen zwischengespeichert und in Textdateien final abgespeichert. Diese Abläufe werden so lange wiederholt, bis alle aus den Langserien extrahierten Sounddateien abgearbeitet und alle zuvor eingeladenen Langserien verarbeitet sind.

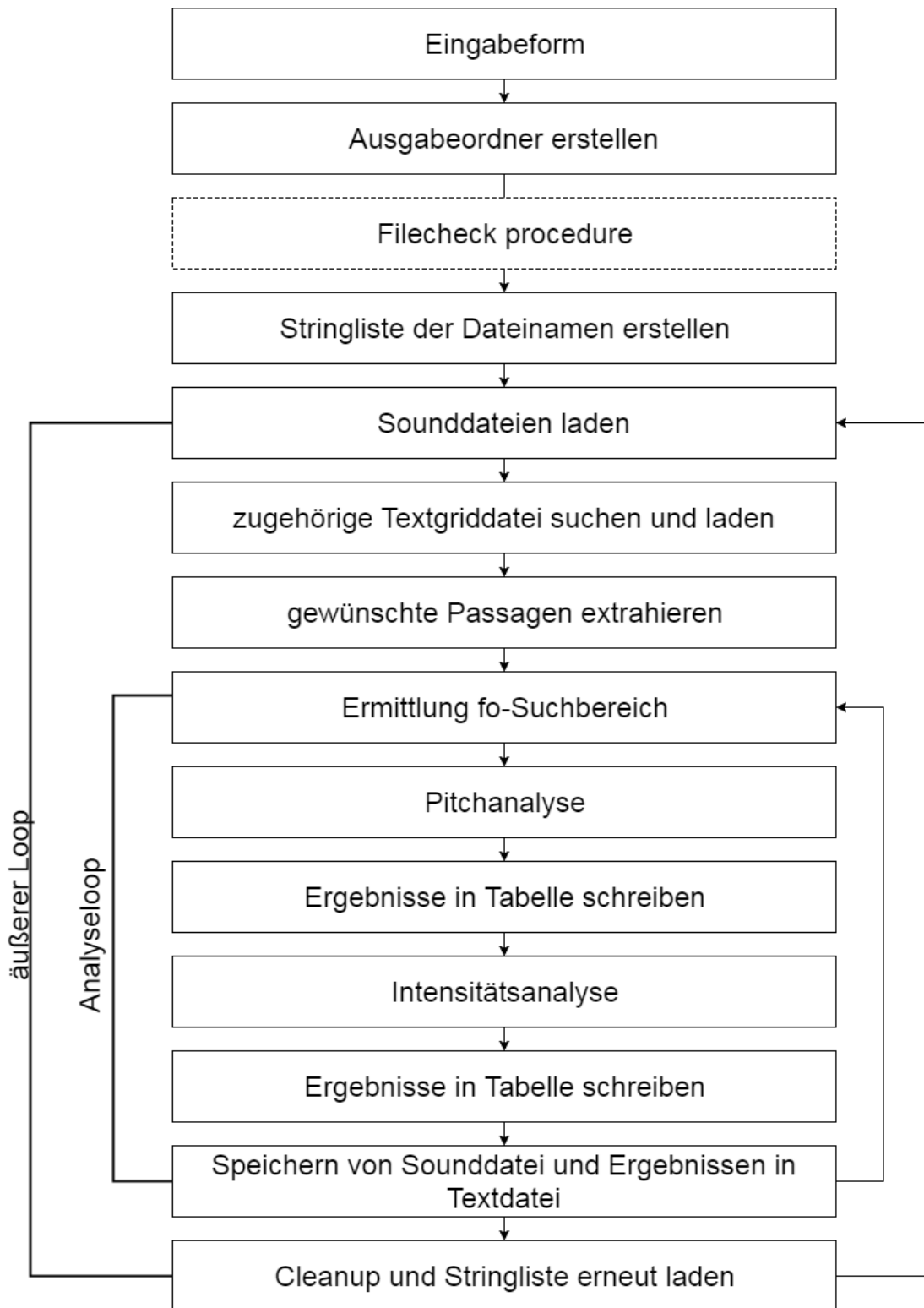


Abbildung 17: Aufbau des erarbeiteten Skriptes

### 3.1.1 Skript zur Analyse annotierter Serien

Nachfolgend werden die einzelnen Abschnitte des Skriptes in Funktion und Aufbau erläutert.<sup>3</sup>

---

<sup>3</sup> Das Kommentieren innerhalb des Skriptes dient dazu, dass der Programmierer selbst im Falle einer Revision, aber auch andere Benutzer oder Programmierer, sich zügig im Skript zurechtfinden können. Kommentare werden durch das Zeichen „#“ am Zeilenanfang gekennzeichnet, welche dem Programm signalisieren, diese Zeilen bei der Ausführung des Skriptes unberücksichtigt zu lassen. Der großzügige Einsatz von Whitespace (Leerzeilen, die nicht bei der Skriptausführung berücksichtigt werden) soll die Lesbarkeit des Skriptes verbessern.



### 3.1.1.1 Eingabeform

Das Skript zur Analyse annotierter Serien beginnt mit dem Codeabschnitt, der zur Erstellung eines grafischen Eingabefensters führt, auf welchem der Benutzer Voreinstellungen und Eingaben vornehmen kann, direkt nachdem das Skript über PRAAT's grafische Oberfläche gestartet wird (Abbildung 18). Beginn und Ende des Codeabschnittes werden durch die beiden Befehle „*form*“ sowie „*endForm*“ gekennzeichnet. Die in Codeabschnitt 1 aufgeführten Befehle, die zur Erstellung des Eingabefensters führen, sind in Anhang A 1.1 detailliert beschrieben.

Run script: Extract sounds, save, write Pitch and Intensity to table.

Directory of sound files.  
C:\Users\Michael\Desktop\  
Directory of TextGrid files.  
C:\Users\Michael\Desktop\  
TextGrid file extension: .TextGrid  
Please specify sounds to extract  
extract intervals: choice  
Output a text file with tier times?  
 timetable  
Pitch File Extension  
output file extension: .tabelle  
Intensity File Extension  
output file extension 2: .tabelle2  
Where should Output folder be created? Be sure to include the final "\\"  
C:\Users\Michael\Desktop\  
Standards Cancel Apply OK

Abbildung 18: Grafisches Eingabefenster

## Codeabschnitt 1:Eingabeform

```
1 #Skript zur automatisierten Verarbeitung von Langserien am ZVES
2 #
3 #Eingabeform mit Dateiauswahl und Voreinstellungen
4 form Extract sounds, save, write Pitch and Intensity to table.
5 comment Directory of sound files.
6 text sound_directory
C:\Users\Michael\Desktop\Wermke\Tests\fertigeanalysetest2
7 comment Directory of TextGrid files.
8 text textGrid_directory
C:\Users\Michael\Desktop\Wermke\Tests\fertigeanalysetest2
9 sentence TextGrid_file_extension .TextGrid
10 comment Please specify sounds to extract
11 optionmenu extract_intervals
12 option e
13 option i
14 comment Output a text file with tier times?
15 boolean timetable 1
16 comment Pitch File Extension
17 sentence output_file_extension .tabelle
18 comment Intensity File Extension
19 sentence output_file_extension_2 .tabelle2
20 comment Where should Output folder be created? Be sure to include
the final "\"
21 text output_directory
C:\Users\Michael\Desktop\Wermke\Tests\fertigeanalysetest2\
22 endform
```

Im Eingabefenster (Abbildung 18) müssen vom Nutzer Dateipfade für die Sounddateien und die Textgriddateien, die die Annotationen enthalten, angegeben werden. Die getrennte Angabe von Dateipfaden für Sound- und Textdateien ermöglicht es, die Dateien aus Ordnungs- oder anderen Gründen in unterschiedlichen Verzeichnissen zu speichern und von dort wieder aufzurufen. Darauf folgt im Fenster die Angabe der Dateiendung für die TextGrid-Dateien.

In den weiteren Eingabefeldern legt der Benutzer nun fest, welche Abschnitte/zeitlichen Intervalle extrahiert werden sollen, ob eine Tabelle mit Zeitwerten der Abschnitte ausgegeben werden soll und schließlich die gewünschten Dateiendungen für die später ausgegebenen Ergebnistabellen und das Verzeichnis, in dem der Ausgabedateienordner erstellt werden soll.

Teile dieser Eingabeform wurden aus einem Skript von Lennes (2003a) übernommen und modifiziert (GNU General Public License).

### 3.1.1.2 Outputordner

Der Eingabeform folgt der Codeabschnitt, der das Dateisystem des Betriebssystems zur Erstellung des Ausgabeordners anweist, in dem die extrahierten Abschnitte und die Analyseergebnisse gespeichert werden sollen (Codeabschnitt 2).

Der Codeabschnitt beginnt zunächst wieder mit einem Kommentar, der die Funktion des Abschnittes kurz erläutert.

Codeabschnitt 2: Outputordner erstellen

```
24 #create Output directory
25 new_output_directory$ = output_directory$ + "output\"
26 createDirectory: new_output_directory$.
```

Nachfolgend wird der gewünschte Dateipfad für den Ausgabeordner in der Stringvariablen „*new\_output\_directory\$*“ gespeichert. Dieser setzt sich aus der zuvor im Eingabefenster gespeicherten Stringvariablen „*output\_directory\$*“ und dem gewünschten Namen des Ausgabeordners, hier „*output*“, zusammen.

Der Befehl „*createDirectory:*“, gefolgt von der zusammengesetzten Stringvariablen als Argument, erstellt nun im zuvor durch den Benutzer festgelegten Verzeichnis einen Ordner mit dem Namen „*output*“.

### 3.1.1.3 Filecheck Procedure

Werden bestimmte Codeabschnitte wiederholt im Skript benötigt, bedient man sich der sogenannten „Procedures“: Der gewünschte Codeabschnitt wird an beliebiger Stelle (jedoch üblicherweise am Anfang oder Ende eines Skriptes) niedergeschrieben und kann dann an beliebiger Stelle des Skriptes wieder aufgerufen („gecallt“) und ausgeführt werden.

Die Procedure im nachfolgenden Codeabschnitt 3 dient an mehreren Stellen im Skript dazu, den Nutzer davor zu schützen, ungewollt Dateien zu überschreiben. Abbildung 19 zeigt das Fenster, das den Benutzer vor ungewolltem Überschreiben einer Datei warnt. Für eine genauere Beschreibung siehe Anhang A 1.2.

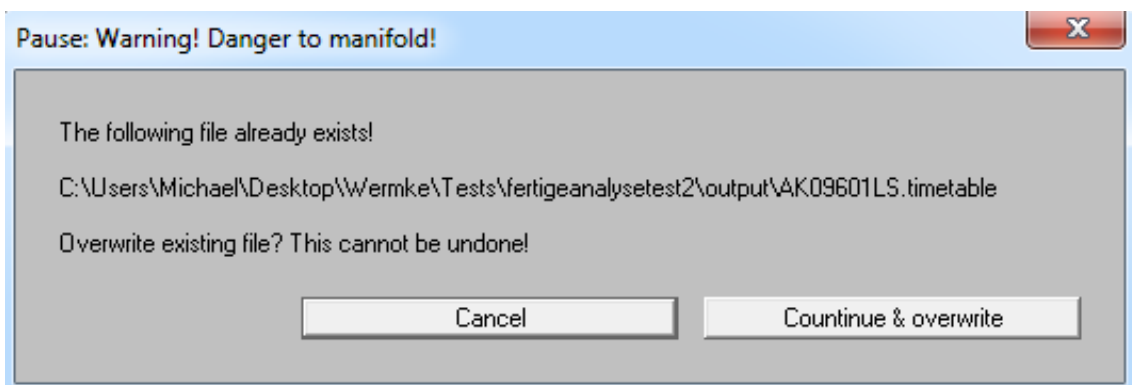


Abbildung 19: Auswahlfenster für Filecheck procedure: Warnung vor versehentlichem Überschreiben der Dateien

### Codeabschnitt 3: Filecheck procedure

```
27
28 #Filecheck Procedure
29 procedure filecheck: .filename$
30 if fileReadable (.filename$)
31 beginPause: "Warning! Danger to manifold!"
32 comment: "The following file already exists!"
33 comment: .filename$
34 comment: "Overwrite existing file? This cannot be undone!"
35 clicked = endPause: "Cancel", "Countinue & overwrite", 1, 1
36 if clicked = 1
37 select all
38 Remove
39 exitScript: "Script was aborted. Please rerun!"
40 endif
41 endif
42 endproc
43
44
```

### 3.1.1.4 Stringliste der Dateinamen erstellen

Um eine alphabetische Liste der Dateinamen im zuvor angegebenen Dateipfad erstellen zu können, welche nachfolgend für die Stapelverarbeitung durch das Skript benötigt wird, nutzt man den Befehl „*Create Strings as file list:*“. Gefolgt wird dieser vom gewünschten Namen des zu erstellenden String-Objekts (in Anführungszeichen) und dem Dateipfad (durch ein Komma getrennt), welcher nach Dateien durchsucht werden soll (vgl. Codeabschnitt 4).

Codeabschnitt 4: Stringliste erstellen

```
44
45 #Stringliste der Dateinamen erstellen
46 strings = Create Strings as file list: "list", sound_directory$ +
"/*.wav"
47 numberOfFiles = Get number of strings
```

Der zuvor in einer Stringvariablen gespeicherte Dateipfad wird hier noch um „*/\*.wav*“ ergänzt. „*\**“ dient in diesem Fall als sogenannte Wildcard. „*/\*.wav*“ legt die Dateiendung der Sounddateien fest, nach denen gesucht werden soll. PRAAT wird vom Skript dadurch angewiesen, den zuvor im Eingabefenster angegebenen Pfad, unabhängig vom Dateinamen, nach allen Dateien mit der Dateiendung „*.wav*“ zu durchsuchen. Durch das der Befehlskette vorgestellte „*strings=*“ wird dem neu entstandenen Strings-Objekt die Variable „*strings*“ zur späteren Verwendung im Skript zugewiesen.

Die Anzahl der durch das Skript gefundenen Dateien wird in der nachfolgenden Programmzeile zunächst durch den Befehl „*Get number of strings*“ abgerufen und in die vorangestellte Variable „*numberOfFiles*“ zur nachfolgenden Verwendung gespeichert.

### 3.1.1.5 Äußerer Loop

Um sukzessive alle zuvor in das Stringsobjekt gespeicherten Sounddateien der Stapelverarbeitung durch das hier erstellte PRAAT-Skript zuzuführen und die gewünschten Analysen abzuarbeiten, greift man auf die Schleifenfunktion der PRAAT-Skriptsprache zurück (Codeabschnitt 5; Anhang A 1.3).

Codeabschnitt 5: Loop durch Stringliste

```
51 for ifile to numberOfFiles
52 filename$ = Get string: ifile
53 Read from file: sound_directory$ + "/" + filename$
54
55 # get the name of the sound object:
56 soundname$ = selected$ ("Sound", 1).
```

Analog zum „*form*“ und „*endform*“ der Eingabeform werden auch Beginn und Ende einer Schleife mit reservierten Begriffen gekennzeichnet.

### 3.1.1.6 Korrespondierendes Textgridfile laden

Damit aus den vorliegenden Langserien nur die für die Analyse relevanten Abschnitte extrahiert werden können, an denen im Anschluss sowohl die Pitch- als auch Intensitätsanalyse erfolgen, ist ein zugehöriges Textgridfile unbedingt notwendig.

Das Skript sucht zur zuvor geöffneten Sounddatei automatisch das passende Textgridfile (Codeabschnitt 6). Die Annotierung der Sounddateien, die in Textgridfiles gespeichert werden, kann zum Beispiel mit dem im Kapitel 3.1.2 beschriebenen Skript erfolgen. Die Namensgebung der Textgridfiles sollte sich am Schema der Sounddateien orientieren, damit diese eindeutig den zugehörigen Tonaufnahmen zugeordnet werden können (für weiterführende Erklärungen siehe Anhang A 1.4).

## Codeabschnitt 6: zugehöriges Textgridfile laden

```
58 # Look for a TextGrid by the same name:
59 gridfile$ = textGrid_directory$ + "/" + soundname$ +
textGrid_file_extension$
60
61 # if a TextGrid exists, open it
62 if fileReadable (gridfile$)
63 Read from file: gridfile$
```

Wurde eine passende Textgriddatei zur zuvor geladenen Sounddatei gefunden und eingeladen, müssen beide Dateien im PRAAT-Objects-Window ausgewählt und somit „aktiv“ geschaltet werden. Dies erfolgt durch den Befehl „*selectObject*:“ gefolgt von Dateityp und -name in Anführungszeichen (Codeabschnitt 7).

## Codeabschnitt 7: Auswählen und aktiv schalten der Sound- und Textgriddatei

```
96
97 #select Soundfile and corresponding Textgrid
98 selectObject: "Sound 'soundname$'", "TextGrid 'soundname$'"
```

Das Verfahren des Ladens der Dateien und passender Textgridfiles orientiert sich in groben Zügen an einem anderen PRAAT-Skript (Lennes, 2003b) (GNU General Public License).



### 3.1.1.7 Codeblock für die automatische Ermittlung temporaler Kenngrößen (Zeitentabelle)

Wurde im Eingabefenster der Haken für die Zeitentabelle gesetzt (vgl. Kapitel 3.1.1.1), wird der Codeblock für die Erstellung einer Tabelle, welche die Zeitwerte der verschiedenen annotierten Abschnitte enthält, ausgeführt (Codeabschnitt 8). Für eine detaillierte Erläuterung siehe Anhang A 1.5.

#### Codeabschnitt 8: Codeblock für Zeitentabelle

```
66 if timetable = 1
67
68 #select Textgrid
69 selectObject: "TextGrid 'soundname$'"
70
71 #Loop für die Name und Zeit der Intervalle
72 numberOfIntervals = Get number of intervals: 1
73 table_time = Create Table with column names: "table", 0, "Name
Startpunkt Endpunkt
Dauer"
74 for intervalNumber from 1 to numberOfIntervals
75 selectObject: "TextGrid 'soundname$'"
76 startTime = Get start point: 1, intervalNumber
77 endTime = Get end point: 1, intervalNumber
78 duration = endTime - startTime
79 name$ = Get label of interval: 1, intervalNumber
80 selectObject: table_time
```

```

81 Append row
82 current_row = Get number of rows
83 Set string value: current_row, "Name", name$
84 Set numeric value: current_row, "Startpunkt", startTime
85 Set numeric value: current_row, "Endpunkt", endTime
86 Set numeric value: current_row, "Dauer", duration
87 endfor
88 timetable_name$ = new_output_directory$ + soundname$ + ".timetable"
89 @filecheck: timetable_name$
90 Save as tab-separated file: timetable_name$
91
92 selectObject: table_time
93 Remove
94
95 endif

```

### 3.1.1.8 Extrahieren der gewünschten Zeitabschnitte

Sind nun eine Sounddatei sowie eine zugehörige Textgriddatei im PRAAT-Objects Window eingeladen und aktiviert, können die gewünschten Passagen automatisch extrahiert werden, um sie weiteren Analysen zuzuführen (Codeabschnitt 9).

#### Codeabschnitt 9: Extraktion gewünschter Passagen

```

100 #Extrahieren der gewünschten Passagen; entspricht dem GUI-Command
101 #Extract intervals where
102 Extract intervals where: 1, "no", "contains", extract_intervals$

```

Die Argumente des Befehls „*Extract intervals where:*“ entsprechen (getrennt durch Kommata) den Eingabefeldern (von oben nach unten) des Eingabefensters „TextGrid & Sound: Extract intervals“ (Abbildung 20). Es kann in der Grafischen Oberfläche im PRAAT-Objects-Window (Abbildung 14) über „Extract -> Extract intervals where...“ aufgerufen werden (Textgrid und Sounddatei müssen dafür angewählt sein).

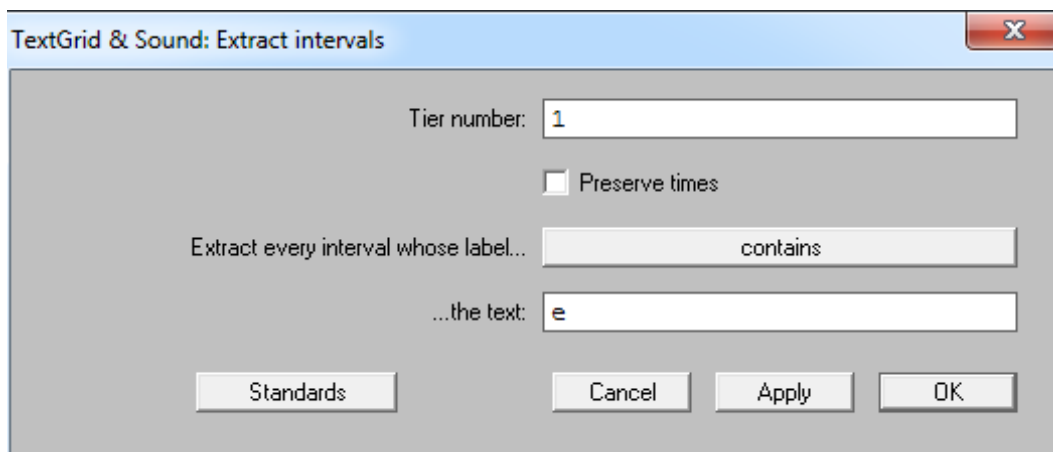


Abbildung 20: Extract intervals GUI-Fenster

In diesem Fall wird Tier 1 des TextGrids nach allen Intervallen abgesucht, die den Buchstaben „e“ enthalten (vgl. Kapitel 1.2.1 Abbildung 2). Diese Intervalle werden anschließend ausgeschnitten und in das PRAAT-Objects-Window als einzelne Soundobjekte ausgegeben. Die soeben ausgegebenen Soundobjekte sind automatisch im Objects-Window markiert (und somit aktiv).

Die Anzahl der neu entstandenen und nun im Objects-Window ausgewählten Soundobjekte wird in die Variable „*numberOfextractedSounds*“ geschrieben. Anschließend werden die Soundobjekte mit einem „inneren“ „*for*“-Loop (Loopzähler „*j*“ ; Maximalanzahl entspricht der in „*numberOfextractedSounds*“ gespeicherten Anzahl) nacheinander ausgewählt und ihr Name jeweils zur späteren Verwendung in einer Variable gespeichert („*sound [j]*“) (Codeabschnitt 10).

## Codeabschnitt 10: Loop durch entstandene Soundfiles

```
104 #Query Anzahl der neuen Files, Anzahl in Variable schreiben
105 numberOfextractedSounds = numberOfSelected ("Sound")
106
107 #Loop2 durch neu entsandenen Objekte im Praat Objects-Window
108 #nacheinander Auswahl aller neu entstandenen Files
109 for j to numberOfextractedSounds
110 sound [j] = selected ("Sound", j)
111 endforGrundfrequenzanalyse (Pitchanalyse)
```

### 3.1.1.9 Grundfrequenz (fo)- und Intensitätsanalyse

Der Erstellung und Speicherung der Pitchfiles (Codeabschnitt 12) ist ein zweistufiges Verfahren vorgeschaltet, das den Suchbereich für die nachfolgende fo-Analyse ermittelt. Hierbei wurde vom Autor ein von Hirst (2011) vorgeschlagenes Voranalyseverfahren verwendet, das, wie von Mayer (2017, pp. 110–112) (creative-commons by-nd-nc) beschrieben, in das Skript implementiert worden ist (Codeabschnitt 11). Das Voranalyseverfahren ist in einem weiteren „inneren“ Loop eingebettet, um so iterativ alle extrahierten Soundobjekte abzuarbeiten. Das Verfahren beinhaltet folgende Schritte: Zunächst wird das aktuell ausgewählte Soundobjekt zusätzlich in eine Stringvariable gespeichert. Ein erstes Pitchobjekt mit festgelegtem Timestep (0,01s) und sehr weitem Pitchrange (fo-Suchbereich) (200-1500Hz) wird erstellt. Anschließend werden obere und untere fo-Quartile berechnet und in Variablen gespeichert ( $q1$  und  $q3$ ). Die so ermittelten Quartile werden anschließend mit empirisch von Hirst (Hirst, 2011) ermittelten Faktoren multipliziert, gerundet und in die Variablen „*floor*“ für die untere und „*ceiling*“ für die obere finale Suchschwelle gespeichert. Das nur temporär benötigte Pitchfile mit weitem Suchbereich wird anschließend per Befehl „*Remove*“ umgehend wieder aus dem Objects-Window entfernt.

## Codeabschnitt 11: Verfahren zur Ermittlung der Suchbereiche „Pitchkorrektur“ nach Mayer (2017)

```
112 #Automatisierte, zweistufige Pitchkorrektur
113 #Pitchanalyse Loop
114 for j to numberOfextractedSounds
115 #Auswählen des Praat-Sound-Objekts
116 selectObject: sound [j]
117 nameofsound$ = selected$ ()
118 #erste Pitchanalyse mit genäherten Werten
119 To Pitch: 0.01, 200, 1500
120 #Q1 ermitteln
121 q1 = Get quantile: 0, 0, 0.25, "Hertz"
122 #Q3 ermitteln
123 q3 = Get quantile: 0, 0, 0.75, "Hertz"
124 #Q1 multiplizieren und runden
125 floor = round (q1 * 0.75)
126 #Q3 multiplizieren und runden
127 ceiling = round (q3 * 2.5)
128 Remove
```

Die durch dieses Verfahren ermittelten Suchbereiche für die obere und untere Grenze der fo-Analyse stehen nun in Variablen gespeichert zur Verfügung. Derzeit wird nur der obere Suchbereich für die fo-Analyse, die sich direkt anschließt, verwendet. Der untere Suchbereich wurde für die aktuelle Anwendung manuell auf 200 Hz festgelegt, es kann jedoch für andere Anwendungen jederzeit stattdessen auf den in der Voranalyse ermittelten Wert zurückgegriffen werden.

Um verarbeitet zu werden, muss das aktuelle Soundobjekt erneut aktiv geschaltet werden. Dies geschieht abermals über den Befehl „*selectObject:*“.

Der Name des Soundobjekts wird wieder aus der Variablen „*sound [j]*“ abgerufen.

In Codeabschnitt 12 sind die Skriptanteile für die fo- sowie Intensitätsanalyse berichtet. Hier werden die selben Funktionen wie im bereits am ZVES vorhandenen Skript von Prochnow verwendet (Prochnow, 2013).

#### Codeabschnitt 12: Pitchanalyse (autocorrelation)

```
131 #Pitchanalyse
132 # Get pitch and intensity tracks
133 selectObject: sound [j]
134 if ceiling <> undefined
135 pitch = To Pitch (ac): 0, 200, 15, "no", 0.03, 0.45, 0.01, 0.35,
0.14, ceiling
136 else
137 pitch = To Pitch (ac): 0, 200, 15, "no", 0.03, 0.45, 0.01, 0.35,
0.14, 600
138 endif
139 selectObject: sound [j]
140 intensity = To Intensity: 200, 0, 1
```

Die Argumente, die dem Befehl „*To Pitch (ac):*“ im Codeabschnitt 12 nachgestellt sind, entsprechen (jeweils durch Kommata getrennt) den Eingabefeldern des Eingabefensters (vgl. Abbildung 21, von oben nach unten), das sich auch manuell über das PRAAT Objects Window aufrufen lässt. Der Timestep, also das Zeitintervall, das zwischen zwei Pitchwerten liegt, die in das Pitchobjekt geschrieben werden, wird automatisch festgelegt und ist abhängig vom Pitchfloor in Hz (Boersma & Weenink, 2019). Dieser wird standardmäßig auf 200 Hz festgelegt, was einen Timestep von 0,00375s ergibt ( $\frac{0,75}{Pitchfloor}$ ), kann aber individuell variiert werden, falls erforderlich (z.B. Komfortlaute im Strohbassregister). Als Pitch-Obergrenze für die fo-Bestimmung wurde der zuvor

im Voranalyserverfahren ermittelte Wert verwendet. Dem neu entstandenen Pitchobjekt wird durch das vorangestellte „*pitch* =“ eine Variable zugeordnet. Kann aus einem Einzelsignal kein Wert für den oberen Schwellenbereich abgeleitet werden (Beispielsweise komplett lautlose Bereiche bei der Analyse inspiratorischer Abschnitte), wird der Text „*undefined*“ statt eines Zahlenwertes in die Variable „*Ceiling*“ gespeichert. Dies führte in Tests zum ungewollten Abbruch des Skriptes. Mit einem Conditional (*if ceiling <> undefined*) wird dies überprüft und stattdessen ein Pitchobjekt mit den von PRAAT vorgegebenen Standardwerten für den Befehl „*To Pitch (ac)* :“ erzeugt (siehe Abbildung 21).

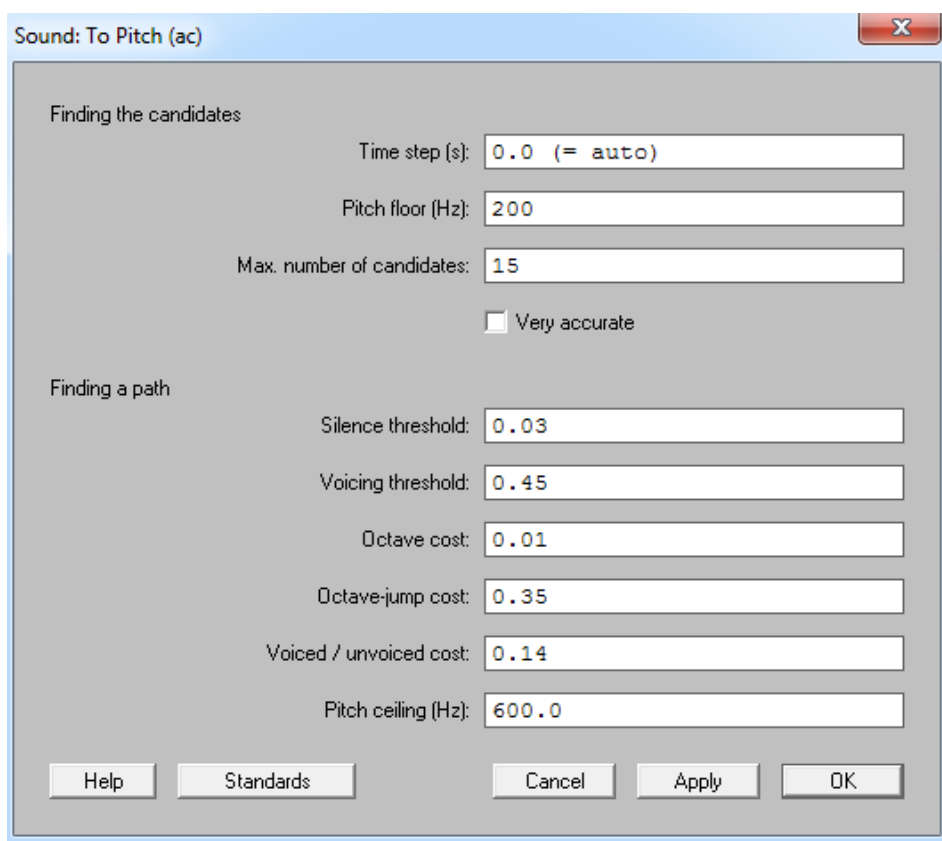


Abbildung 21: Sound to pitch (ac) Eingabefenster

Nun wird das Soundobjekt erneut ausgewählt, um ein Intensitätsobjekt zu erzeugen. Dies erfolgt mit dem Befehl „*To Intensity*: “: Der Pitchfloor wurde manuell auf 200 Hz festgelegt, die restlichen Argumente wurden auf den von PRAAT für diesen Befehl vorgegeben Standardwerten belassen. Wie zuvor beim Pitchobjekt wird auch dem Intensitätsobjekt wieder eine Variable zugewiesen,

hier schlicht „*intensity*“. Für die nachfolgende Melodieanalyse im CDAP wird der jeweils verwendete Timestep und die jeweiligen Suchgrenzen mit übergeben. Die restlichen Werte wurden auf den am ZVES verwendeten Standardeinstellungen belassen (basierend auf einem Skript von Prochnow (2013) und empirischer Testung).

### 3.1.1.10 Framespezifische Zeitwerte in Variablen speichern

Die Zeitwerte (frames) werden gemeinsam mit den Pitch- und Intensitätswerten in Textdateien gespeichert und können zur Weiterverarbeitung an ZVES-interne Systeme (oder beliebige andere Systeme) übergeben werden. (Codeabschnitt 13).

Codeabschnitt 13: Zeiten in Variablen schreiben

```
142 #Get times
143 selectObject: intensity
144 numberOfFramesIntensity = Get number of frames
145 selectObject: pitch
146 numberOfFramesPitch = Get number of frames
147 writeInfoLine: numberOfFramesIntensity
148 appendInfoLine: numberOfFramesPitch
```

Dazu werden zunächst die Pitch- respektive Intensitätsobjekte im PRAAT-Info-Window abermals durch „*selectObject*“ aktiv geschaltet, sodann wird die Gesamtanzahl an Frames sowohl für das erzeugte Pitchobjekt als auch für das erzeugte Intensitätsobjekt, jeweils in eine Variable abgelegt („*numberOfFramesIntensity*“ und „*numberOfframesPitch*“).

Das Verfahren, mit dem die Ausgabe in Textdateien erfolgt, orientiert sich ebenfalls an einem von A. Prochnow verwendeten Skript (Prochnow, 2013) jedoch mit dem entscheidenden Unterschied, dass die einzelnen Werte zunächst mit der in PRAAT integrierten Tabellenfunktion in eine Tabelle geschrieben und dann komplett gespeichert werden.



In parallel zur Skriptearbeitung laufenden Tests zeigte sich, dass das bisherige Vorgehen wie bei Prochnow (ibid.), die Werte zunächst im PRAAT Info Window auszugeben und dann per Befehl „*appendFileLine*“ aus dem Infofenster in eine Textdatei zu kopieren, mit größeren Datenmengen, sprich längeren Sounddateien, überlastet war und zu häufigen Abstürzen von PRAAT führte.

Der Flaschenhals scheint hierbei der Schreibvorgang und Dateizugriff außerhalb der PRAAT-Umgebung über das Dateisystem des (z.B. Windows- oder Unix-) Betriebssystems zu sein. Wird nun der komplette Datensatz zunächst programmintern in einer Tabelle gespeichert und dann komplett an eine Textdatei übergeben, trägt dies nicht nur zur Stabilität bei sondern verkürzt auch den Gesamtvorgang entscheidend.

### 3.1.1.11 Werte in Tabelle schreiben und anschließend Textdateien speichern

Zuerst werden zwei Tabellenobjekte erstellt „*Create Table with column names:*“. Den Objekten werden, wie üblich, Variablen zugewiesen, um später wieder einfach auf sie zugreifen zu können (*table*, *table2*). Die Argumente sind abermals durch Kommata getrennt und entsprechen dem Eingabefenster der grafischen Oberfläche. (Tabellenname, Anzahl der Zeilen, in Anführungszeichen Name der Spalten) (vgl. Codeabschnitt 14).

Codeabschnitt 14: Tabellenobjekte erstellen

```
151 #in Tabelle schreiben
152 table = Create Table with column names: "table", 0, "time pitch"
153 table2 = Create Table with column names: "table", 0, "time
intensity"
```

Mithilfe eines weiteren Loops (Zähler *iframe*, Maximalanzahl aus *numberOfFramesPitch*) wird nun für jedes Frame des Pitchobjekts ein Zeitwert per Befehl „*Get time from frame number:*“ und ein dazugehöriger Pitchwert in Hertz mit dem Befehl „*Get value in frame number:*“ in jeweils

eine Variable geschrieben (Codeabschnitt 15). Eine detaillierte Erklärung findet sich in Anhang A 1.6.

#### Codeabschnitt 15: Werte in Tabelle schreiben (Pitch)

```
155 # Iterate over the pitch and intensity tracks, one sample at a
time
156 for iframe to numberOfFramesPitch
157 selectObject: pitch
158 time = Get time from frame number: iframe
159 selectObject: pitch
160 pitchVal = Get value in frame: iframe, "Hertz"
161
162 selectObject: table
163 Append row
164 current_row = Get number of rows
165 Set numeric value: current_row, "time", time
166 Set numeric value: current_row, "pitch", pitchVal
167 endfor
```

Das Vorgehen zum Ausgeben und Abspeichern der Werte für die Intensität verhält sich analog: In einem weiteren Loop (Codeabschnitt 16) wird zunächst die Gesamtanzahl an Frames eines (Pitch- oder Intensitäts-) Objektes in einer Variablen gespeichert und anschließend die zugehörigen Werte Frame für Frame in Variablen zwischengespeichert. Nach dem Auswählen des zweiten Tabellenobjektes (*table2*) werden die Werte dann in ebendieses wieder abgelegt. Auch hier werden die Werte Zeile für Zeile in eine neue Tabelle geschrieben.

## Codeabschnitt 16: Werte in Tabelle schreiben (Intensität)

```
169 # Iterate over intensity tracks
170 for jframe to numberOfFramesIntensity
171   selectObject: intensity
172   timeint = Get time from frame number: jframe
173   selectObject: intensity
174   intensityVal = Get value in frame: jframe
175
176   selectObject: table2
177   Append row
178   current_row = Get number of rows
179   Set numeric value: current_row, "time", timeint
180   Set numeric value: current_row, "intensity", intensityVal
181 endfor
```

Die Implementierung im Skript orientierte sich in groben Zügen an einem geposteten Skript im PRAAT User Forum (Mahrt, 2016).

Abschließend erfolgt die Speicherung (Codeabschnitt 17) aller in den Tabellen gespeicherten Pitch-, Intensitäts- und den zugehörigen Zeitwerten in Textdateien und die Entfernung nicht mehr benötigter Objekte (Codeabschnitt 18). Für eine detailliertere Beschreibung wird auf Anhang A 1.7 verwiesen.

## Codeabschnitt 17: Speichern der Dateien

```
182 #Tabelle als neues Textfile speichern (filereadable testen!)
183 selectObject: table
184 nameofpitchfile$ = new_output_directory$ + nameofsound$ +
output_file_extension$
185 @filecheck: nameofpitchfile$
186 endif
187 Save as tab-separated file: nameofpitchfile$
188 if ceiling <> undefined
189 appendFileLine: nameofpitchfile$, "Settings:" + " " + nameofsound$
+
"Floor" + " " + string$ (floor) + "Ceiling" + " " + string$ (ceiling)
190 else
191 appendFileLine: nameofpitchfile$, "Standard settings because
Ceiling
undefined"
192 endif
193 selectObject: table2
194 nameofintensityfile$ = new_output_directory$ + nameofsound$ +
output_file_extension_2$
195 @filecheck: nameofintensityfile$
196 Save as tab-separated file: nameofintensityfile$
197 appendFileLine: nameofintensityfile$, "Settings:" + " " +
nameofsound$ +
"Floor" + " " + string$ (floor) + "Ceiling" + " " + string$ (ceiling)
198 selectObject: sound [j]
199 nameofwavfile$ = new_output_directory$ + nameofsound$ + ".wav"
200 @filecheck: nameofwavfile$
201 Save as WAV file: nameofwavfile$
```

## Codeabschnitt 18: Cleanup

```
202 # Cleanup
203 selectObject: pitch
204 Remove
205
206
207 selectObject: intensity
208 Remove
209
210 selectObject: table
211 Remove
212
213 selectObject: table2
214 Remove
215
216 selectObject: sound [j]
217 Remove
218
219 endfor
220
221
222
```

Sind alle extrahierten Soundobjekte eines Textgrid- und Langserienpaares schließlich abgearbeitet, werden auch diese per „*selectObject*“ zunächst aktiv geschaltet und dann entfernt (Codeabschnitt 19). Nachfolgend wird die Stringliste wieder aktiv geschaltet und das nächste Textgrid- und Langserienpaar in das Objektwindow geladen und die oben genannten Analysen erneut durchgeführt. Sind keine Sounddateien mehr in der Stringliste vorhanden, wird auch der „äußere“ Loop per „*endfor*“ abgeschlossen und das Skript terminiert an dieser Stelle.

#### Codeabschnitt 19: Cleanup großer Loop

```
223 #cleanup
224 selectObject: "Sound 'soundname$'", "TextGrid 'soundname$'"
225 Remove
226
227 #Stringliste erneut auswählen
228 selectObject: "Strings list"
229
230
231 #Loop2ende
232 endfor
233
234 selectObject: strings
235 Remove
```

## 3.1.2 Skript zur halbautomatischen Annotation von Langserien

### 3.1.2.1 Übersicht

Obgleich das Hauptziel der Arbeit in der Erstellung eines Skriptes zur automatischen fo- und Intensitätsanalyse von Einzellauten innerhalb einer längeren Aufnahmesequenz (Langserie) in dem Kooperationsprojekt (Kapitel 2.1) lag, musste im Zuge der Entwicklung ein weiteres Skript erstellt werden. Es dient dazu, die erforderliche vorbereitende Annotation der Langserien zu vereinfachen und dadurch eine erhebliche Arbeitszeiterparnis und verbesserte Reproduzierbarkeit für das aktuelle, vor allem aber geplante Folgeprojekt mit größerer Probandenanzahl zu erreichen. Die automatische Identifikation und anschließende Annotation in einer Textgriddatei in „laute“ und „lautlose“ Anteile einer Langserie wurde auf Basis ihrer Intensitätskurve realisiert. Das Skript liefert einen automatisch erstellten Vorschlag zur Aufteilung einer Lautsequenz (Langserie), die dann im weiteren Verlauf manuell durch einen Benutzer angepasst und abgespeichert werden kann. Auch hier soll der komplette Funktionsumfang kurz in Form eines Fließdiagrammes verdeutlicht werden (Abbildung 22). In den folgenden Unterkapiteln werden die Funktionen kurz erläutert; für ausführlichere Beschreibungen des Textcodes wird wiederum jeweils auf die entsprechenden Abschnitte des Anhangs verwiesen.

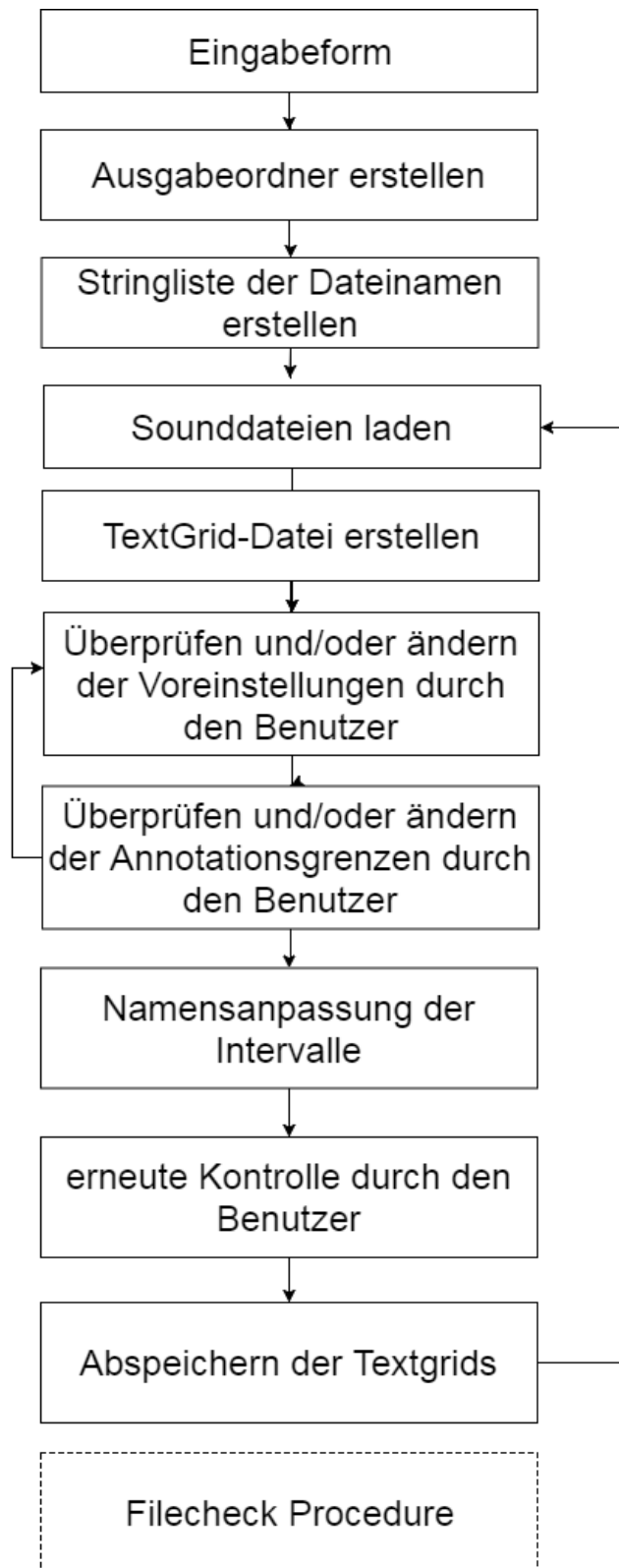


Abbildung 22: Übersicht über den Programmumfang des Skriptes zur halbautomatischen Annotation von Langserien



### 3.1.2.2 Skriptbeschreibung

Dieses Skript beginnt mit einem Kommentar (Code Annotation 1), gekennzeichnet durch „#“, der in Kürze Funktion und Titel erläutert. Weiterhin ist ein kurzer Verweis auf den Entwickler und eine Kontaktmöglichkeit angegeben. Das Skript wird nach Abschluss der Dissertation online frei zur Verfügung gestellt.

#### Code Annotation 1: Kommentar

```
1 #Michael's Little Annotationhelper v3 Beta
2 #
3 #Script to semi-automatically annotate sounding and silent parts
4 #of a sound.
5 #Lets the user check and change the settings and annotation and
6 #saves the TextGrid files automatically.
7 #
8 # © 2018 Michael Eisenmann
9 # for Questions and licensing contact
10 # mail@michaelsimoneisenmann.de
```

### 3.1.2.3 Eingabeform

Dem initialen Kommentar folgt eine Eingabeform (Code Annotation 2) die, wie in Kapitel 3.1.1.1 beschrieben, Eingaben durch den Benutzer in einer grafischen Oberfläche zulässt. In diesem Skript werden hier lediglich die Dateipfade für die zu annotierenden Sounddateien sowie den durch das Skript erstellten Ausgabeordner vom Nutzer eingefügt.

#### Code Annotation 2: Eingabeform

```
12 #Input form for file paths
13
14 form Michael's Little Annotation helper
15 comment Directory of sound files.
16 text sound_directory
C:\Users\Michael\Desktop\Wermke\Annotationhelper_Demo
17 sentence textGrid_file_extension .TextGrid
18 comment Where should Output folder be created? Be sure to include
the final "\"
19 text output_directory
C:\Users\Michael\Desktop\Wermke\Annotationhelper_Demo\
20 endform
```

### 3.1.2.4 Ausgabeordner und Stringliste der Dateinamen erstellen

Analog zu Kapitel 3.1.1.2 lässt man auch hier einen Ausgabeordner für die Annotationsdateien erstellen (Code Annotation 3).

Code Annotation 3: Ausgabeordner

```
22 #create output directory
23
24 new_output_directory$ = output_directory$ + "output\"
25 createDirectory: new_output_directory$
26
```

Wie in Kapitel 3.1.1.4 beschrieben, wird auch für dieses Skript eine Dateiliste der Sounddateien benötigt, die mit den folgenden Codezeilen erstellt wird (Code Annotation 4).

Code Annotation 4: Stringliste für Dateinamen erstellen

```
27 #Search for soundfiles and write the names into a string object;
write number
28 #of sound files into a variable
29
30 strings = Create Strings as file list: "list", sound_directory$ +
"/*.wav"
31 numberOfFiles = Get number of strings
```

### 3.1.2.5 Sounddateien laden und Textgriddateien erstellen

Nachfolgend beginnt der Loop, der nun nacheinander die Sounddateien aus der Stringliste in PRAAT einlädt (Code Annotation 5). Zur näheren Erläuterung wird auf Kapitel 3.1.1.5 und Anhang A 1.3 verwiesen.

Code Annotation 5: Laden der Sounddateien

```
32
33 #Open sound files in the string object consecutively
34
35 for ifile to numberOfFiles
36 selectObject: strings
37 filename$ = Get string: ifile
38 Read from file: sound_directory$ + "/" + filename$
39 soundname$ = selected$ ("Sound", 1)
```

Um dem Benutzer einen ersten Annotationsvorschlag zu liefern, wird auf Basis der Intensitätskurve des zu annotierenden Lautes ein Textgrid-Objekt erstellt, das ein Annotationstier mit „stillen“ und „lauten“ Intervallen enthält. Hierzu wird auf die Funktion „*To Textgrid (silences)*“ zurückgegriffen. Die durch Kommata getrennten Argumente legen den Intensitätsschwellenwert (-20 dB) für die Annotation „stiller“ Abschnitte, Minimal- und Maximaldauer „stiller“ bzw. Laute enthaltender Abschnitte (0,1s) und schließlich deren Label fest (können geändert werden). Als Label wurden „i“ für Stille und „e“ für Laute enthaltende Abschnitte vergeben. Dem so neu angelegten Textgrid-Objekt wird zur späteren Verwendung die Variable „textgrid“ zugewiesen (Code Annotation 6).

Code Annotation 6: Erstellen der Textgridobjekte

```
40 #Create TextGrid file
41 textgrid = To TextGrid (silences): 100, 0, -20, 0.1, 0.1, "i", "e"
```

### 3.1.2.6 Überprüfen und/oder ändern der Voreinstellungen durch den Benutzer

Das Skript schreitet mit einer erstmaligen Überprüfung des Annotationsvorschlages durch den Benutzer fort (Code Annotation 7).

Der Befehl „*beginPause:*“ mit nachfolgendem Kopfzeilentitel „*Does it come close?*“ in Anführungszeichen pausiert das Skript und öffnet, gleichzeitig mit einer Editorumgebung für die aktuelle Sounddatei, ein Auswahlfenster (Abbildung 23). Es ermöglicht dem Benutzer, die voreingestellten Werte zu überschreiben und eine Textgriddatei mit anderer unterer fo-Suchgrenze (Beschriftung im Fenster: *pitch*) sowie anderem Intensitätsschwellenwert (Beschriftung im Fenster: *threshold*) zu erzeugen; gesetzt den Fall, dass das mit Standardwerten automatisch erzeugte TextGrid noch zu weit von einer korrekten Annotation abweicht. Die Änderung von Intensitätsschwellenwert oder fo-Suchbereich führen in der Regel zu einer besseren Annäherung an eine, im Sinne der Zielsetzung, korrekte Annotation; eine manuelle Nachannotation ist nur noch in Einzelfällen bei sehr komplexen Einzellaute und/oder schlechtem Signal-Rausch-Abstand erforderlich. Für eine detailliertere Beschreibung des Codes siehe auch Anhang A 1.8.

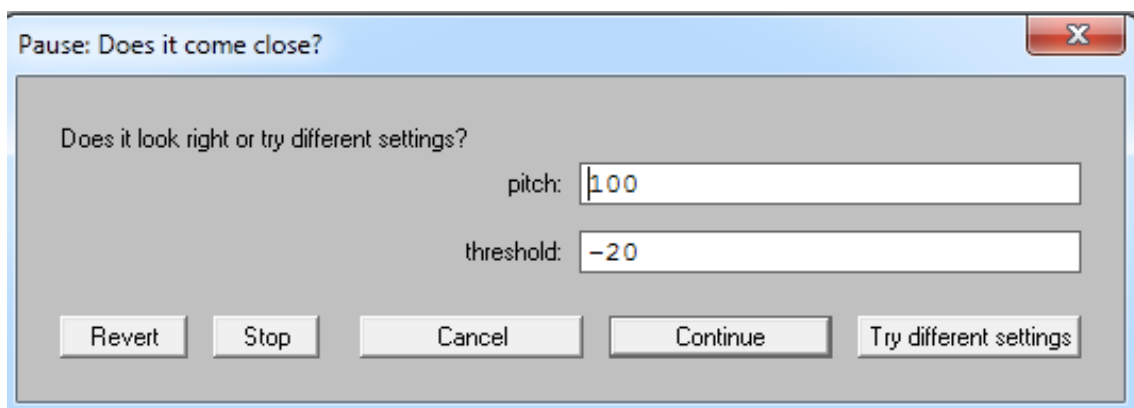


Abbildung 23: Auswahlfenster Voreinstellungen (fo-Suchgrenze (*pitch*) in Hz, Intensitätsschwellenwert (*threshold*) in dB)

Code Annotation 7: Codeabschnitt für Dialogfenster zur manuellen Überprüfung durch Benutzer

```
42 #Let the user check the automatic annotation settings and change at
will
43 repeat
44 beginPause: "Does it come close?"
45 selectObject: "Sound 'soundname$', "TextGrid 'soundname$'"
46 View & Edit
47 comment: "Does it look right or try different settings?"
48 real: "pitch", "100"
49 real: "threshold", "-20"
50 clicked = endPause: "Cancel", "Continue", "Try different settings",
2
51 if clicked = 1
52 removeObject: "Sound 'soundname$', "TextGrid 'soundname$'"
53 exitScript: "Script aborted, please rerun!"
54 elsif clicked = 3
55 selectObject: "TextGrid 'soundname$'"
56 Remove
57 selectObject: "Sound 'soundname$'"
58 textgrid = To TextGrid (silences): pitch, 0, threshold, 0.1, 0.1,
"i", "e"
59 endif
60 until clicked = 2
```

Das Skript enthält eine zweite Pause zur Überprüfung der Annotation: Ist der Benutzer mit dem Vorschlag durch die automatische Routine zufrieden, erhält er nochmals die Gelegenheit, die Annotation final zu überprüfen und gegebenenfalls die Grenzen erneut anzupassen, indem abermals per „*beginPause*“ die Ausführung des Skriptes kurz unterbrochen wird. Klickt der Benutzer im Pausendialogfenster auf „*Continue*“, wird auch diese Pause beendet. Es folgt der Codeblock für die fortlaufende Nummerierung der jeweils annotierten Abschnitte (vgl. Code Annotation 8).

Code Annotation 8:           Codeabschnitt für Dialogfenster zur erneuten Überprüfung des Annotationsergebnisses

```
61 #second Pause for checking the boundaries of the annotation
62 beginPause: "Check the annotation and continue when done!"
63 endPause: "Continue", 1
64 endeditor
```

### 3.1.2.7 Nummerierung der Intervalle

Um die Unterscheidung der Einzellaute in weiteren Analyseschritten zu erleichtern und ihre sichere, eindeutige Zuordnung zu gewährleisten, erfolgt eine aufsteigende Nummerierung der „stillen“ beziehungsweise Laute enthaltenden Abschnitte (siehe Code Annotation 9 und Anhang A 1.9). Die Umbenennung der Intervalle erfolgt angelehnt an einen Ansatz in einem Skript zur Silbenerkennung von Mayer (2012).

Code Annotation 9: Codeabschnitt zur Namensanpassung der Intervalle

```
65 #automatically rename the sounding and nonsounding intervals with
increasing
numbers
66 selectObject: "TextGrid 'soundname$'"
67 numberOfIntervals = Get number of intervals: 1
68 schrei_no = 0
69 leise_no = 0
70 for ints to numberOfIntervals
71 select textgrid
72 label$ = Get label of interval: 1, ints
73 if label$ = "e"
74 schrei_no = schrei_no+1
75 Set interval text: 1, ints, "e" + string$ (schrei_no)
76 elsif label$ = "i"
77 leise_no = leise_no+1
78 Set interval text: 1, ints, "i" + string$ (leise_no)
79 endif
```



### 3.1.2.8 Abspeichern der Textgrids

Das Ende des Skriptes bildet das Speichern der Textgrid-Dateien (vgl. Kapitel 3.1.1.11). Dieser Codeblock ist in Code Annotation 10 zu sehen.

Code Annotation 10: Codeabschnitt zur Speicherung der Textgrids

```
81 #Save the checked and renamed TextGridFiles
82 textgridname$ = new_output_directory$ + soundname$ +
textGrid_file_extension$
83 @filecheck: textgridname$
84 Save as text file: textgridname$
85 selectObject: "Sound 'soundname$'", "TextGrid 'soundname$'"
86 Remove
87 endfor
```

Zunächst wird der gewünschte Name des Textgrids in der Stringvariablen aus Outputdateipfad, Soundname und Dateiendung des Textgridfiles zusammengesetzt. Anschließend wird die Filecheck Procedure aufgerufen, um ein versehentliches Überschreiben der Datei zu vermeiden. Die Sounddatei und das soeben abgespeicherte TextGrid werden in der Objektliste zunächst per „*selectObject*“ ausgewählt und anschließend per Befehl „*Remove*“ entfernt. Der große Loop terminiert schließlich per „*endfor*“.

Dem eigentlichen Codeblock des Skriptes nachgestellt befindet sich auch an dieser Stelle eine Filecheck Procedure (Code Annotation 11). Diese Procedure kann aus jedem Teil des Skriptes aufgerufen werden und dient dazu, den Nutzer vor versehentlichem Überschreiben bereits zuvor erstellter Dateien zu schützen (vgl. Kapitel 3.1.1.3 und Anhang A 1.2).

## Code Annotation 11: Filecheck Procedure

```
91 #Filecheck Procedure
92 procedure filecheck: .filename$
93 if fileReadable (.filename$)
94 beginPause: "Warning! File already exists!"
95 comment: "The following file already exists!"
96 comment: .filename$
97 comment: "Overwrite existing file? This cannot be undone!"
98 clicked = endPause: "Cancel", "Countinue & overwrite", 1, 1
99 if clicked = 1
100 select all
101 Remove
102 exitScript: "Script was aborted. Please rerun!"
103 endif
104 endif
105 endproc
```

## 3.2 Teil 2: Ergebnisse der Skripttestung

### 3.2.1 Ergebnisse der teilautomatischen Annotation

Nachfolgend werden die Ergebnisse der mithilfe des Skriptes erfolgten (Kapitel 3.1.2) automatisierten Annotation der Testsignale sowie deren gegebenenfalls notwendige manuelle Nachbearbeitung durch einen Benutzer berichtet (teilautomatisches Verfahren). Alle Cursorpositionen, an denen Fehler oder Artefakte bei der teilautomatisierten Annotation beobachtet wurden und die manuelle Korrekturen notwendig gemacht haben, werden hier aufgezeigt. Die Diskussion erfolgt in Kapitel 4. In der Diskussion wird zusammenfassend auf Besonderheiten, wie beispielsweise Segmentierungspausen in den Säuglingslauten, die häufig Annotationsfehler verursachen, eingegangen. Zur Beschreibung der PRAAT-Diagramme siehe Kapitel 2.4 im Methodenteil.

#### 3.2.1.1 Ergebnisse der skriptbasierten Annotation der Testsignale aus der Patientengruppe der Probanden mit LKGS

##### **Testsignal A:**

Abbildung 24 zeigt die Ergebnisse der automatischen Annotation für das Testsignal A aus einer Spontanlotaufnahme eines Säuglings mit Gaumenspalte aus der Gruppe der Probanden mit LKGS-Spaltbildungen (Kapitel 2.3.1.1). In Abbildung 25 ist die finale Annotation nach der manuellen Korrektur durch einen Benutzer dargestellt.

Ein Vergleich der Cursorpositionen in beiden Abbildungen zeigt, dass die automatische Annotation durch das Skript bei diesem Testsignal, mit Ausnahme des vierten Lautes, das gewünschte Ergebnis geliefert hat. Der automatisch erstellte Vorschlag des Skriptes hat sowohl die ersten beiden Laute mit einbögiger Melodie (blaue Kurven) als auch den dritten Weinaut mit komplexerer Struktur korrekt annotiert. Auch die lautlosen Bereiche (i) wurden korrekt erkannt. Beim vierten Laut (e4) musste aufgrund spezifischer Signaleigenschaften nachträglich eine manuelle Anpassung (Cursorpositionsverschiebung) durch den Benutzer vorgenommen werden: Hier handelt es sich bei dem „stillen“ Bereich zwischen den Einzelbögen um eine Segmentierungspause innerhalb des Lautes

(Abbildung 25). Dieser Abschnitt wurde zwar vom Skript korrekt als lautloser Anteil (i) erkannt, inhaltlich jedoch handelt es sich um eine kurze oszillatorische Pause der Stimmlippen (Segmentierung durch laryngeale Konstriktion). Zwei Grenzen („Boundaries“) wurden manuell entfernt. Auch die Cursorposition, die das Ende des Lautes markiert, wurde nach hinten verschoben. Der Weinaut enthält eine zweite, längere Segmentierungspause und endet mit einem sehr kurzen phonierten Element, bevor die zweite Inspiration folgt.

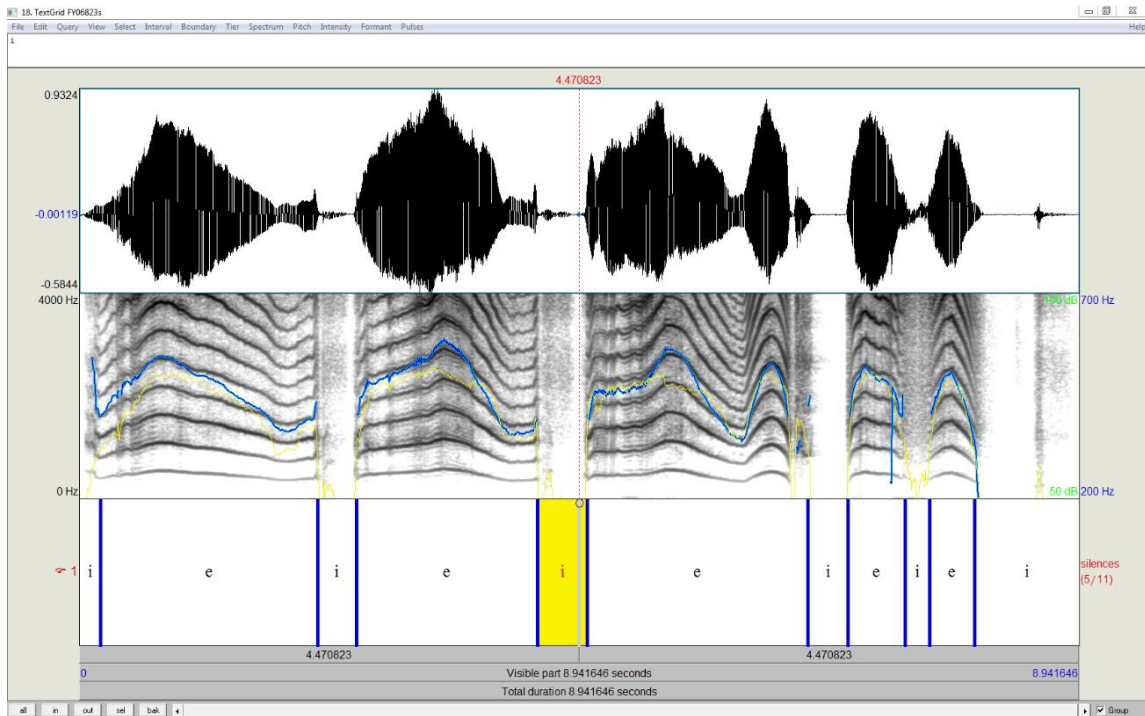


Abbildung 24: Testsignal A: automatisch generierter Annotationsvorschlag von oben nach unten: Zeitbereichsdarstellung, Schmalbandspektrogramm und Annotationstier mit Vorschlägen für laute (e) und „stille“ (i) Anteile eines Säuglings mit LKGS im Alter von 68 Tagen

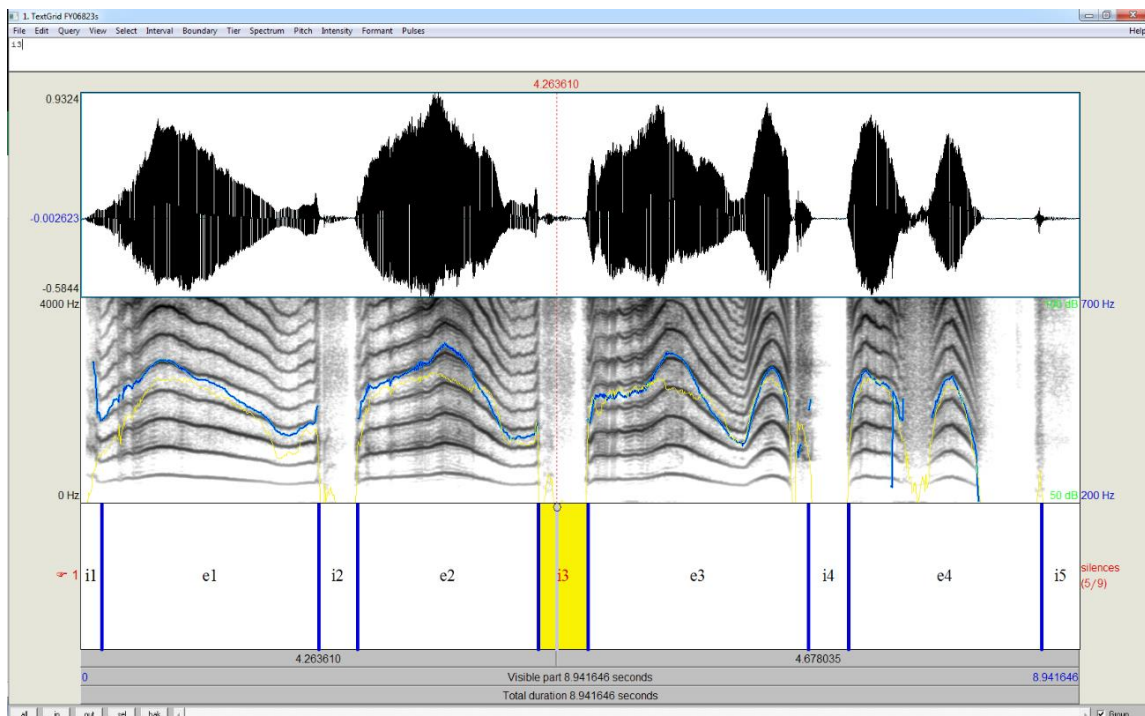


Abbildung 25: Testsignal A: manuell korrigierte Annotation von oben nach unten: Zeitbereichsdarstellung, Schmalbandspektrogramm und angepasste sowie nummerierte Annotationen für laute (e) und lautlose (i) Anteile eines Säuglings mit LKGS im Alter von 68 Tagen

## Testsignal B:

Abbildung 26 zeigt das Ergebnis der automatischen Annotation durch das Skript für das Testsignal B (Kapitel 2.3.1.2). Nach einem ersten Skriptdurchlauf zeigte sich eine Fehlannotation des dritten Weinlautes. Daraufhin wurde eine zweite automatische Annotation mit verändertem Intensitätsschwellenwert durchgeführt (vgl. Kapitel 3.1.2.6). Das Ergebnis ist in Abbildung 27 dargestellt.

Aufgrund der Hintergrundgeräusche und Lauteigenschaften wurden manuelle Anpassungen vorgenommen. Abbildung 27 zeigt die Annotation nach manueller Anpassung des Schwellenwertes (vgl. Kapitel 3.1.2.6), welche die Erkennung der korrekten Lautgrenzen bereits erheblich verbessert. Beim ersten „i“ in Abbildung 26 handelt es sich allein um Hintergrundrauschen, dies wurde in der finalen Version (Abbildung 28) manuell mit „noise“ gekennzeichnet. Die Cursor für die Lautgrenzen der Laute e1, e2 und e3 wurden geringfügig verschoben, um einen weiteren Lautbereich an den Rändern mit einzubeziehen. Beim Laut e3 wurden zwei Annotationsgrenzen entfernt, da es sich hier nur um eine Segmentierungspause handelte (vgl. Abbildung 27 und Abbildung 28).

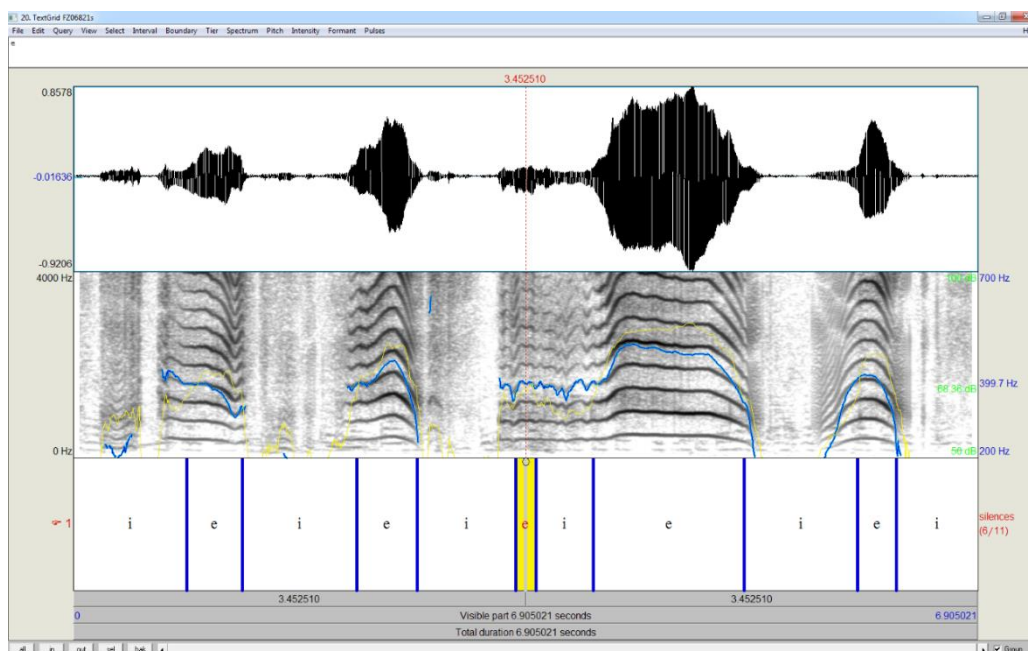


Abbildung 26: Testsignal B: automatisch generierter Annotationsvorschlag von oben nach unten: Zeitbereichsdarstellung, Schmalbandspektrogramm und Annotationstier mit Vorschlägen für laute (e) und lautlose (i) Anteile eines Kindes mit LKGS im Alter von 68 Tagen mit eingesetzter Oberkieferplatte

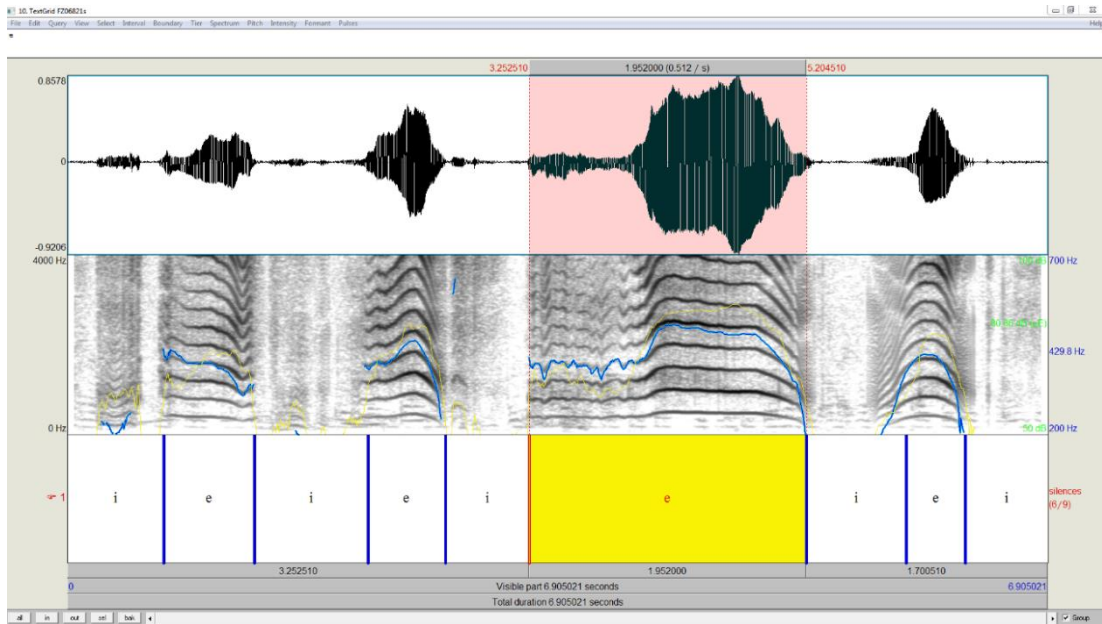


Abbildung 27: Testsignal B: automatisch generierter Annotationsvorschlag nach Anpassung des Schwellenwertes  
 von oben nach unten: Zeitbereichsdarstellung, Schmalbandspektrogramm und Annotationstier mit Vorschlägen für laute (e) und lautlose (i) Anteile eines Kindes mit LKGS im Alter von 68 Tagen mit eingesetzter Oberkieferplatte

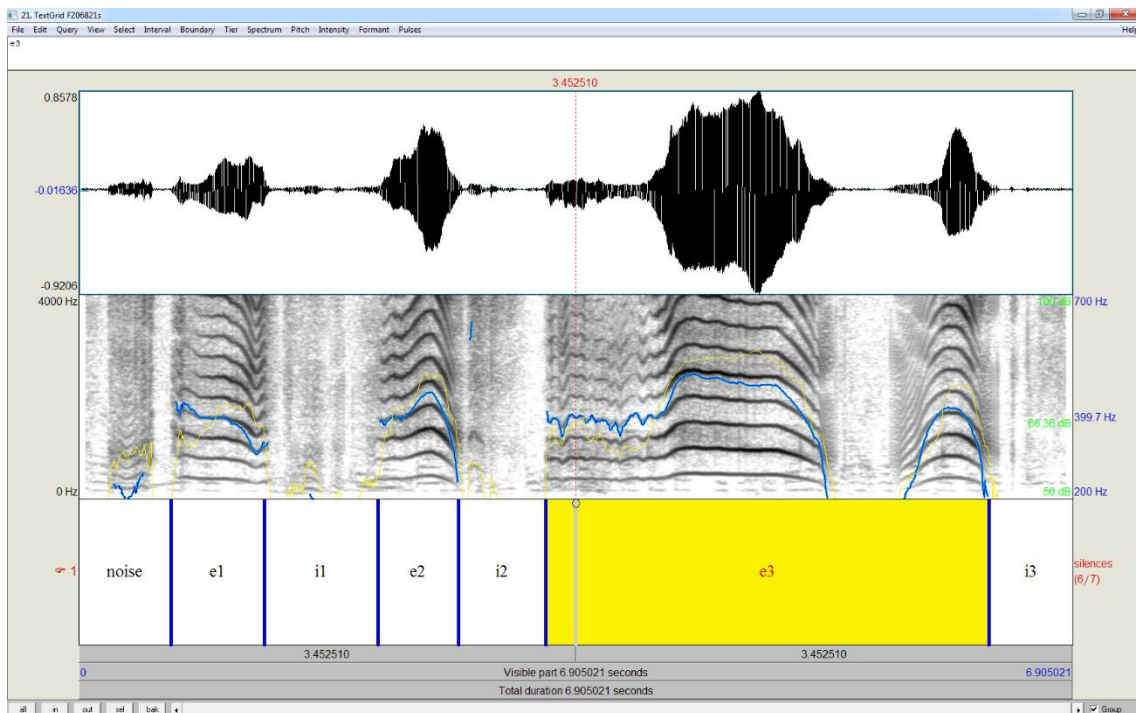


Abbildung 28: Testsignal B nach manueller Korrektur  
 von oben nach unten: Zeitbereichsdarstellung, Schmalbandspektrogramm und angepasste sowie nummerierte Annotationen für laute (e) und lautlose (i) Anteile eines Säuglings mit LKGS im Alter von 68 Tagen mit eingesetzter Oberkieferplatte

### **Testsignale C und D:**

Die Testsignale C und D enthielten keine Wein- sondern spontane Sprachlaute (Drei- und Vierwortsatz).

Für eine detaillierte Beschreibung der Signale wird auf Kapitel 2.3.1.3 verwiesen. Im letzten Wort wurde ein „i“ annotiert, hier handelt es sich um eine Zwischensilbenpause im Wort (vgl. Abbildung 30).

Die automatische Annotation lieferte für Testsignal C ein relativ gutes Ergebnis. Das Wort „Opa“ wurde trotz Segmentierung bei dessen Wiederholung korrekt erkannt. Eine manuelle Cursorverschiebung musste lediglich nach dem Wort „Bäume“ vorgenommen werden (vgl. Abbildung 29 und Abbildung 30).

Beim finalen Annotat wurden die gesprochenen Wörter im Tier annotiert (Abbildung 30). Im vorderen Teil wurde noch eine inspiratorische Pause ergänzt. Im hinteren Teil wurden die Grenzen der beiden letzten Wörter korrigiert.

Im Unterschied zu Testsignal C lieferte die automatische Annotation des Testsignals D ein, vergleichsweise, schlechtes Ergebnis (Abbildung 31). Es handelt sich um dasselbe Kind im Alter von 926 Tagen (vgl. Kapitel 2.3.1.3). Der gesamte vordere Teil der Sequenz musste manuell nachbearbeitet werden. Nur die Grenzen von i4, e4 und i5 (Abbildung 32) wurden entsprechend des automatischen Vorschlages beibehalten.



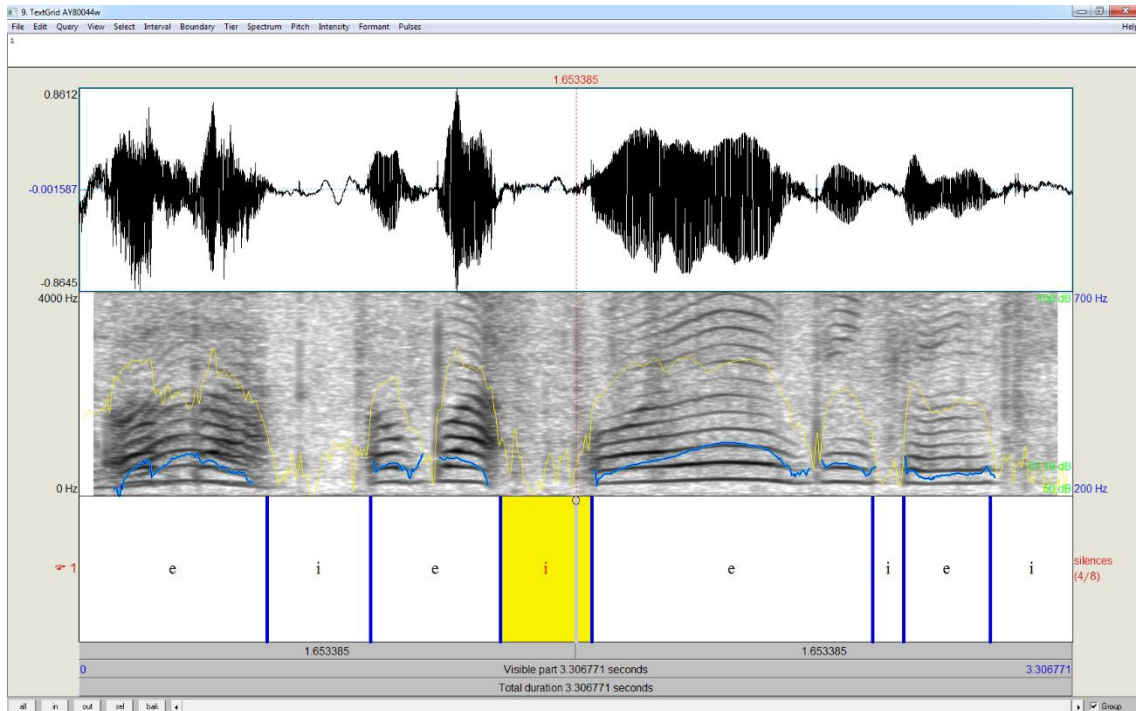


Abbildung 29: Testsignal C: automatisch generierter Annotationsvorschlag von oben nach unten: Zeitbereichsdarstellung, Schmalbandspektrogramm und Annotationstier mit Vorschlägen für laute (e) und lautlose (i) Anteile eines Kindes mit LKGS, 800 Tage

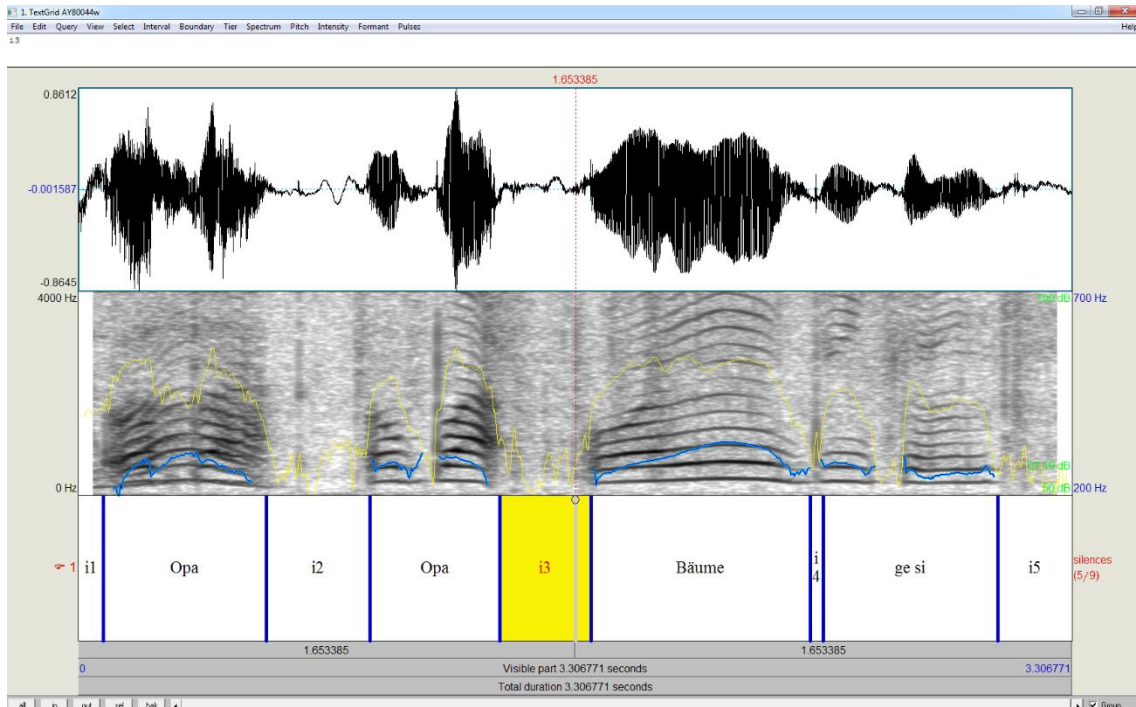


Abbildung 30: Testsignal C mit manuell annotierten Wörtern von oben nach unten: Zeitbereichsdarstellung, Schmalbandspektrogramm und angepasste Annotationen für Wörter und lautlose (i) Anteile eines Kindes mit LKGS im Alter von 800 Tagen

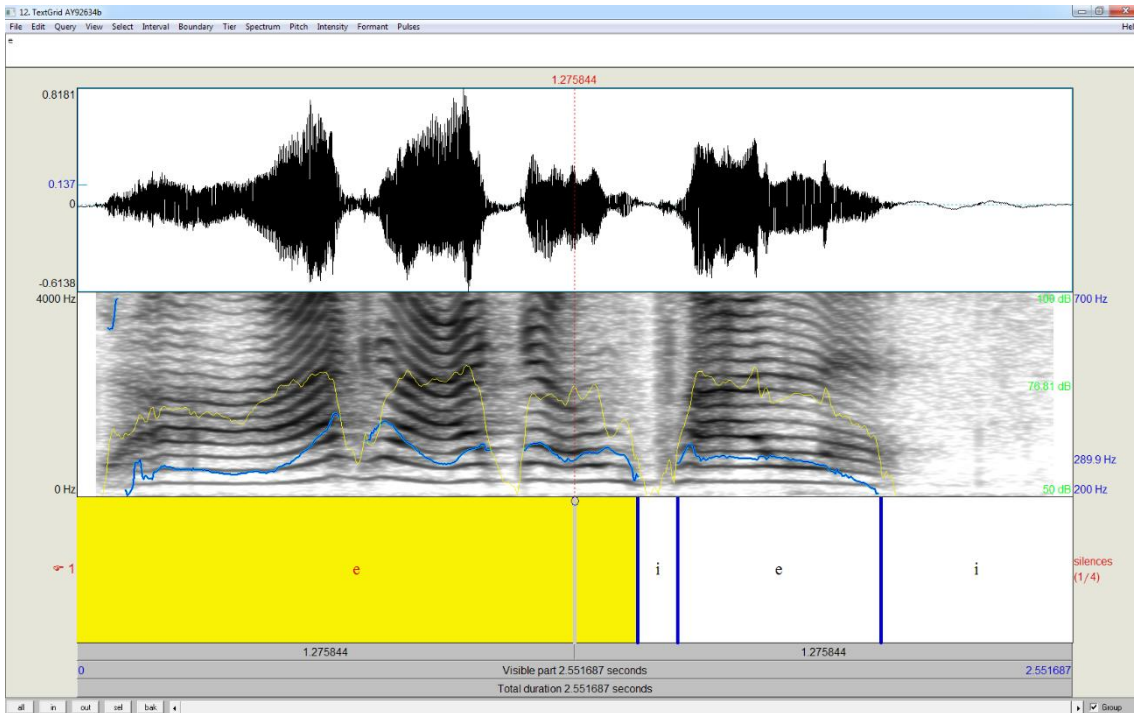


Abbildung 31: Testsignal D: automatisch generierter Annotationsvorschlag von oben nach unten: Zeitbereichsdarstellung, Schmalbandspektrogramm und Annotationstier mit Vorschlägen für laute (e) und lautlose (i) Anteile eines Kindes mit LKGS, 926 Tage

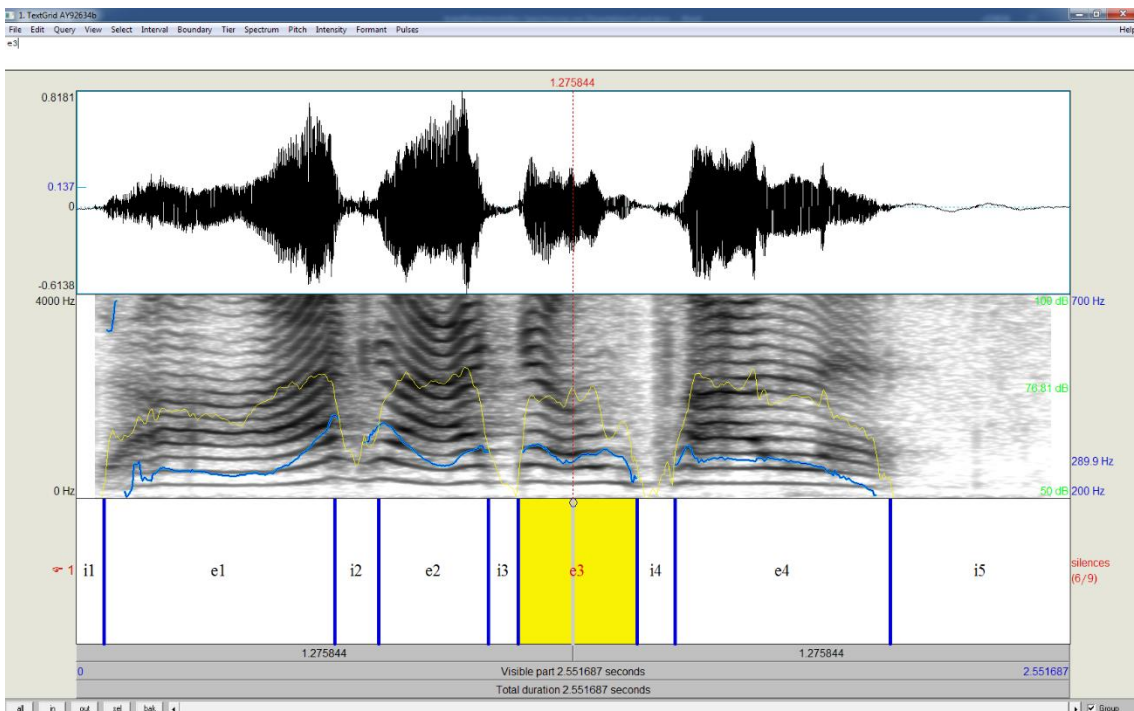


Abbildung 32: Testsignal D nach manueller Korrektur der Annotation von oben nach unten: Zeitbereichsdarstellung, Schmalbandspektrogramm und angepasste sowie nummerierte Annotationen für laute (e) und lautlose (i) Anteile eines Säuglings mit LKGS im Alter von 926 Tagen

### 3.2.1.2 Ergebnisse der skriptbasierten Annotation der Testsignale aus der Patientengruppe der Säuglinge mit Sagittalnahtsynostosen

#### Testsignale E, F und G:

Im Testsignal E werden die Laute, mit Ausnahme des letzten, mit geschlossenem Mund produziert (Nasallaute „hm“). Dadurch hat  $f_0$  die höchste Energiedichte aller Harmonischen. Die automatische Erstellung eines Vorschlags für die Annotation, die sich auf die Hüllkurvenintensität bezieht, hat hier gut funktioniert (vgl. Abbildung 33 und Abbildung 34): Es waren im Prinzip keine manuellen Anpassungen notwendig. Lediglich die Grenze des Lautes e5 (Abbildung 34) wurde final leicht nach hinten verschoben, um hier das Ausschwingphänomen (sogenannter „Endlaut“) mit zu erfassen.

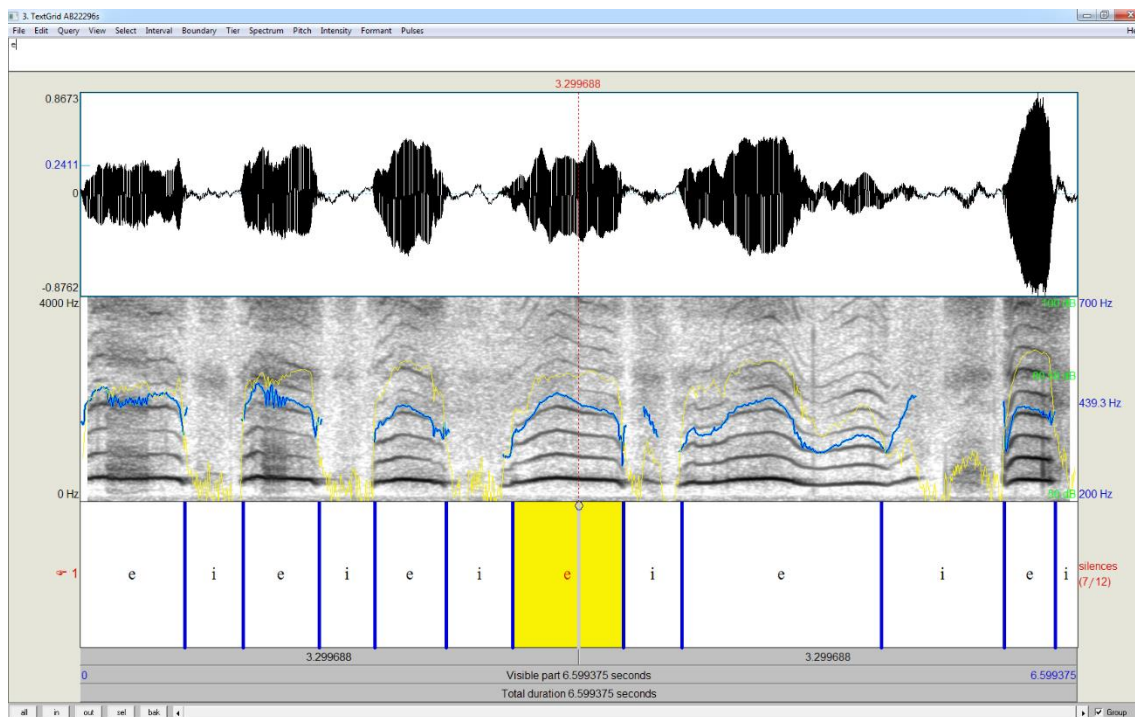


Abbildung 33: Testsignal E: automatisch generierter Annotationsvorschlag von oben nach unten: Zeitbereichsdarstellung, Schmalbandspektrogramm und Annotationstier mit Vorschlägen für laute (e) und lautlose (i) Anteile eines Säuglings mit Sagittalnahtsynostose, 222 Tage

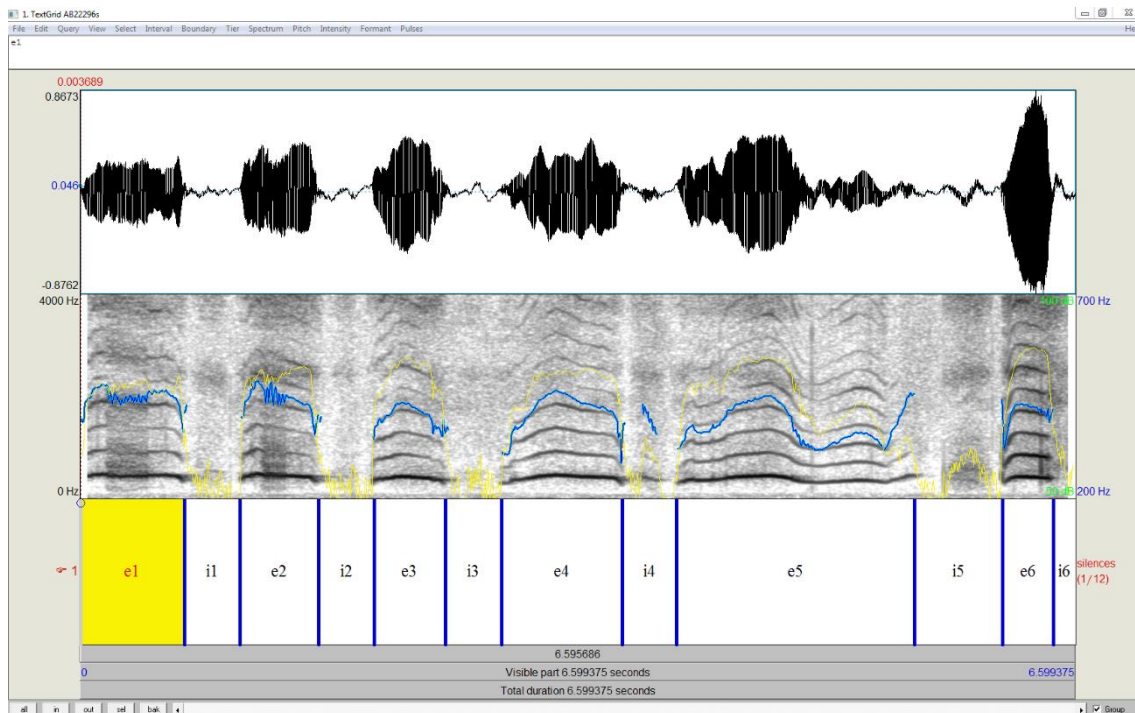


Abbildung 34: Testsignal E: nach manueller Korrektur der Annotation von oben nach unten: Zeitbereichsdarstellung, Schmalbandspektrogramm und angepasste sowie nummerierte Annotationen für laute (e) und lautlose (i) Anteile eines Kindes mit Sagittalnahtsynostose, 222 Tage

Abbildung 35 zeigt das automatisch generierte Annotationsergebnis durch das Skript für Testsignal F (für detaillierte Beschreibung siehe Kapitel 2.3.1.4).

In diesem Beispiel waren die Inspirationsgeräusche zwischen den einbögigen, einfachen Babbellauten sehr prominent. Aufgrund der dennoch deutlichen Intensitätsunterschiede zwischen expiratorischen Lauten und inspiratorischen Geräuschen, lieferte die automatische Annotation ein korrektes Ergebnis, welches unverändert in das finale Annotat übernommen werden konnte. Die Babbellaute und inspiratorischen Anteile wurden durch das Skript aufsteigend durchnummeriert und so zur Weiterverarbeitung vorbereitet (Abbildung 36).

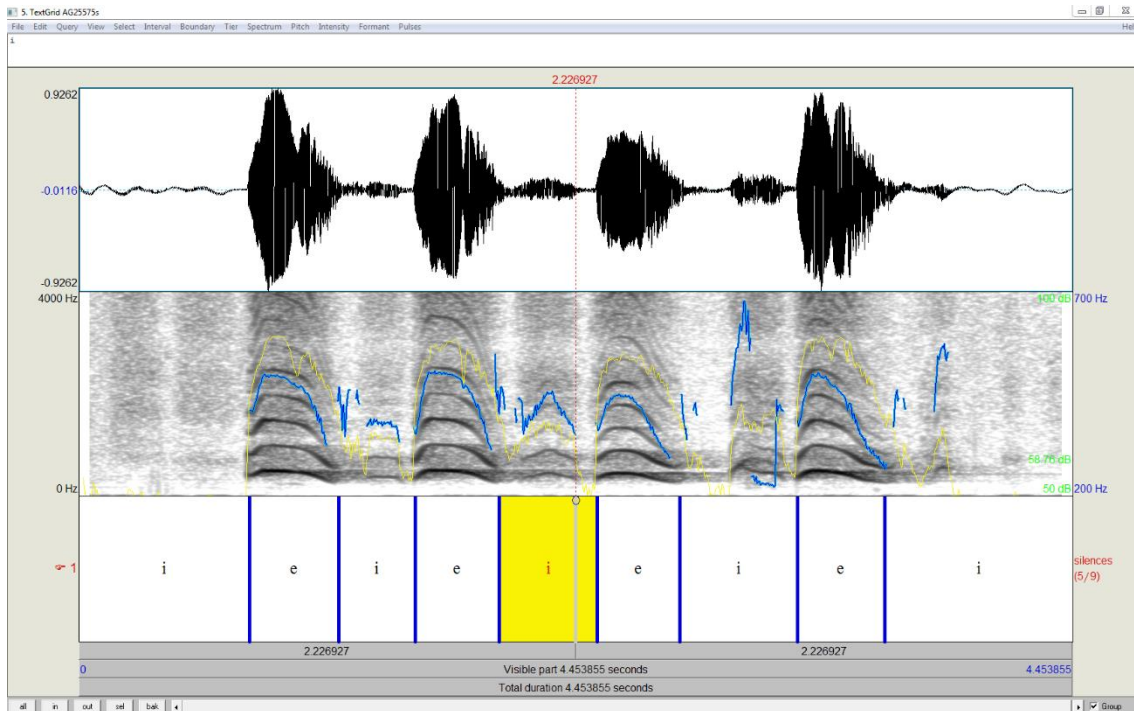


Abbildung 35: Testsignal F: automatisch generierter Annotationsvorschlag von oben nach unten: Zeitbereichsdarstellung, Schmalbandspektrogramm und Annotationstier mit Vorschlägen für laute (e) und lautlose (i) Anteile eines Kindes mit Sagittalnahtsynostose, 255 Tage

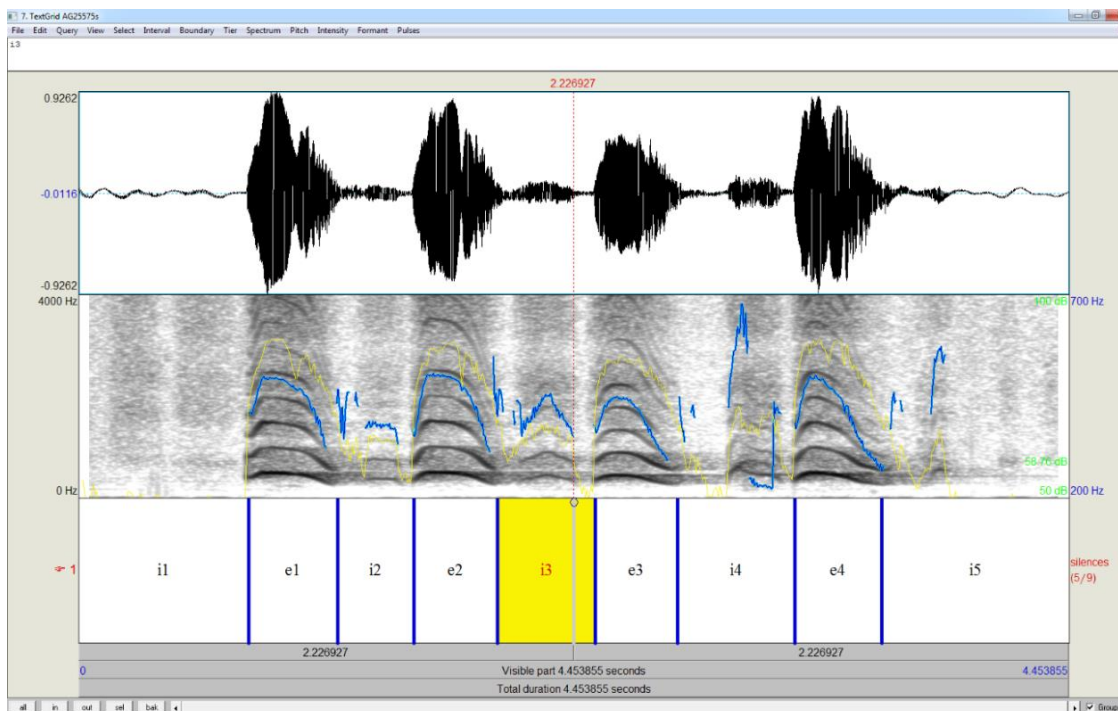


Abbildung 36: Testsignal F nach automatischer Nummerierung der Label von oben nach unten: Zeitbereichsdarstellung, Schmalbandspektrogramm und angepasste sowie nummerierte Annotationen für laute (e) und lautlose (i) Anteile eines Kindes mit Sagittalnahtsynostose, 255 Tage

Testsignal G enthält vier kanonische Babbellaute, die im Kopfregeister artikuliert wurden. Auch das Inspirationsgeräusch nach der ersten Doppelsilbe ist sehr hochfrequent und hat eine annähernd gleich hohe Intensität wie der vorhergehende Babbellaut (vgl. Zeitdarstellung in Abbildung 37). Das ist auch der Grund dafür, dass die Inspirationsphase „i2“ (Abbildung 38) im Annotationsvorschlag des Skriptes fälschlicherweise als expiratorische Phase („e“) annotiert wurde (drittes „e“ in Abbildung 37). Eine analoge Fehlannotation trat im zweiten doppelsilbigen Babbellaut auf: Auch hier wurde der innerlautliche Zwischensilbenbereich im automatisch generierten Vorschlag als „i“, also Zwischenlautpause annotiert (vgl. „e2“ in Abbildung 38).

Aufgrund dieser besonderen Signaleigenschaften mussten daher bei den ersten zwei Lauten e1 und e2 in Abbildung 38, entgegen der Vorschläge des automatisch generierten Annotationsvorschlages, Annotationsgrenzen entfernt und die Lautgrenzen angepasst werden. Die Vorschläge im hinteren Teil des Testsignals waren korrekt und wurden unverändert zur weiteren Analyse übernommen.

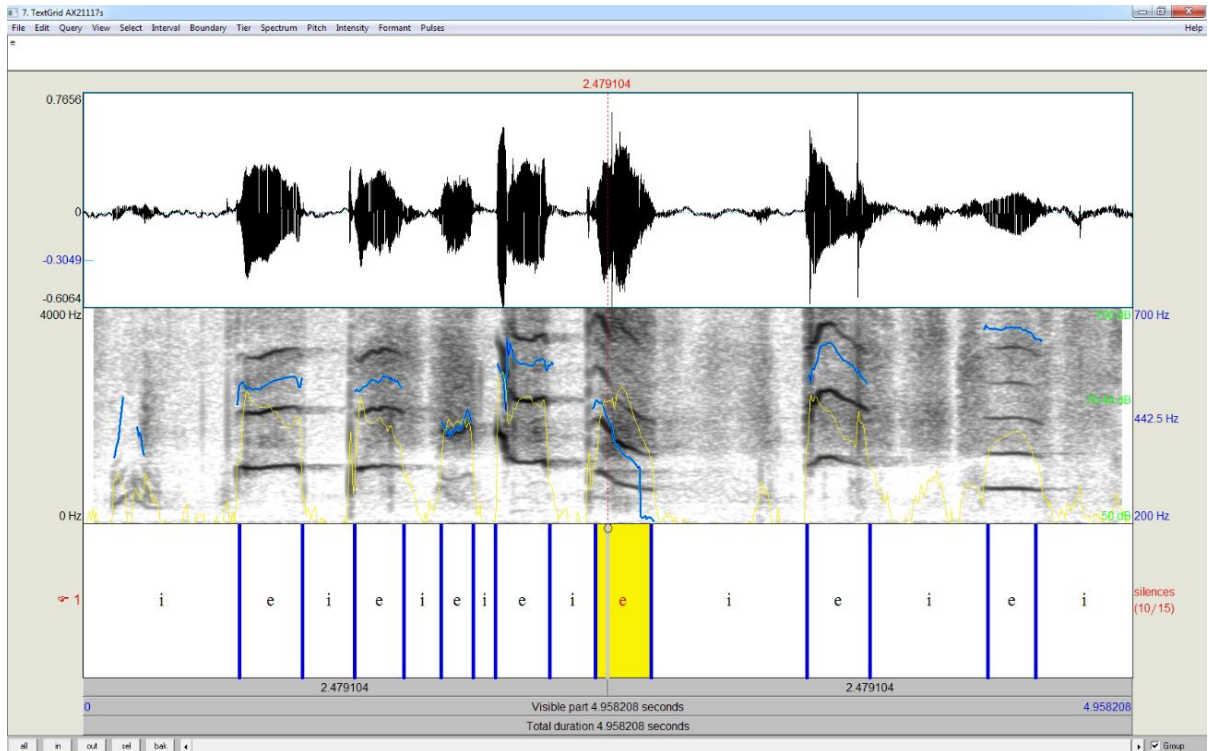


Abbildung 37: Testsignal G: automatisch generierter Annotationsvorschlag von oben nach unten: Zeitbereichsdarstellung, Schmalbandspektrogramm und Annotationstier mit Vorschlägen für laute (e) und lautlose (i) Anteile eines Säuglings innerhalb einer Kontrollgruppe, 211 Tage

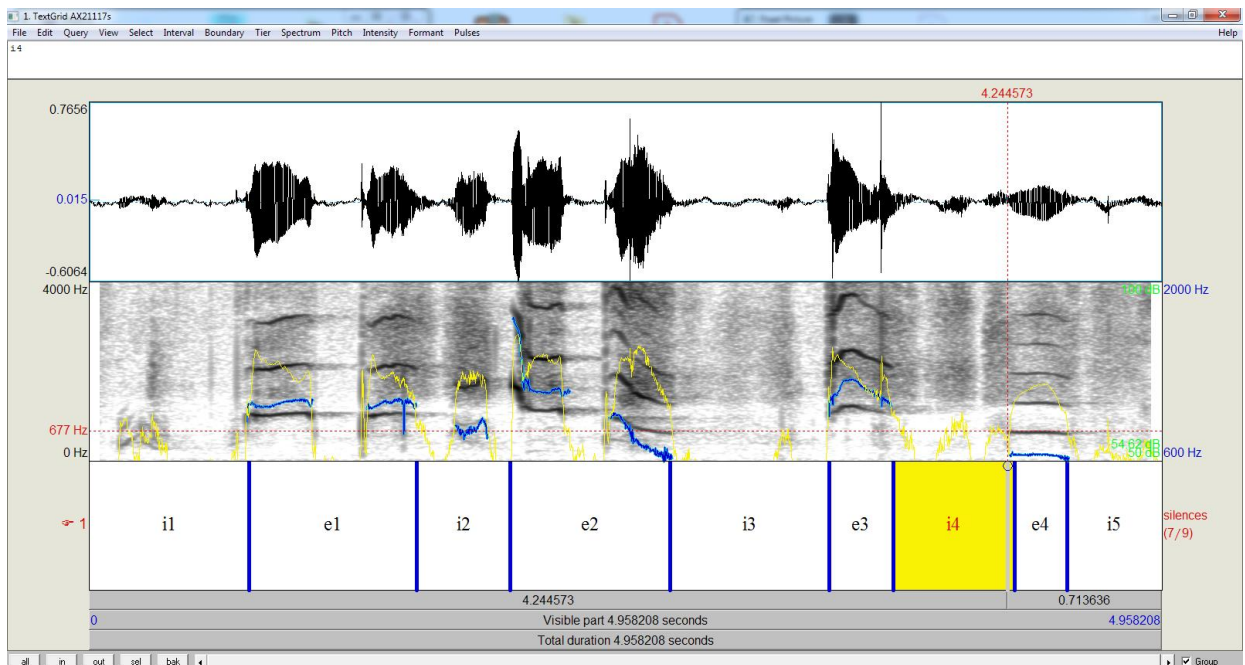


Abbildung 38: Testsignal G nach manueller Korrektur der Annotation von oben nach unten: Zeitbereichsdarstellung, Schmalbandspektrogramm und angepasste sowie nummerierte Annotationen für laute (e) und lautlose (i) Anteile eines Kindes innerhalb einer Kontrollgruppe, 211 Tage

### 3.2.1.3 Ergebnisse der skriptbasierten Annotation der Testsignale aus der Patientengruppe der Säuglinge mit sensori-neuraler Hörbeeinträchtigung

Die automatisch generierten Annotationsergebnisse in Abbildung 39 stammen von **Testsignal H** (Kapitel 2.3.3). Obwohl es sich hier um eine vergleichsweise lange und komplex rhythmisierte Sequenz spontaner Weinlaute handelt, lieferte die skriptbasierte automatische Annotation sehr gute Ergebnisse. Lediglich der erste spezifisch rhythmisierte Laut e1 (vgl. Abbildungen Abbildung 39 und Abbildung 40) wurde formal zwar richtig, inhaltlich aber inkorrekt annotiert (vgl. Diskussion in Kapitel 4.1.2.1). Da es sich im vorderen Teil bei diesem Beispiel inhaltlich um Segmentierungspausen handelt (e1 Abbildung 40), wurden hier Lautgrenzen im Vergleich zum Vorschlag in Abbildung 39 angepasst. Im hinteren Teil der Sequenz wurden die Lautgrenzen entsprechend dem Vorschlag des Skriptes unverändert übernommen und durch das Skript entsprechend aufsteigend durchnummeriert.

Das automatische Annotationsergebnis für **Testsignal I** (Kapitel 2.3.3) ist in Abbildung 41 berichtet.

Ähnlich wie bei Testsignal H wurden manuell bei den ersten beiden Lauten e1 und e2 (Abbildung 42) Cursor in der Mitte der Laute aus inhaltlichen Gründen entfernt, da hier abermals Segmentierungspausen im Vorschlag des Skriptes (Abbildung 41) als inspiratorische Bereiche annotiert wurden.



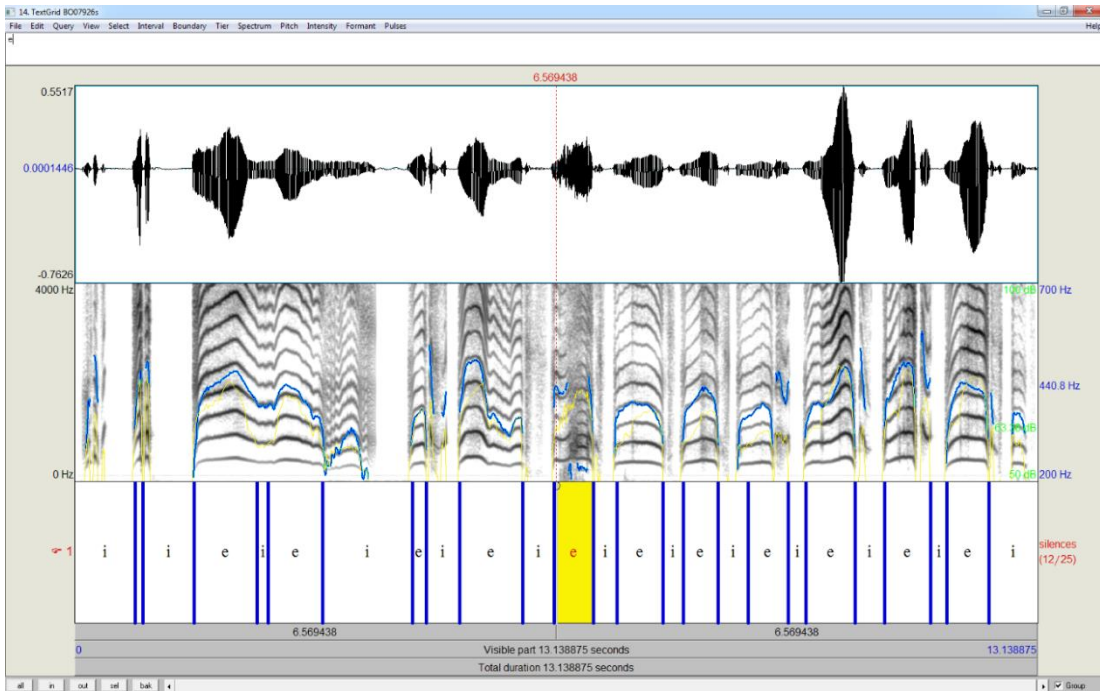


Abbildung 39: Testsignal H: automatisch generierter Annotationsvorschlag von oben nach unten: Zeitbereichsdarstellung, Schmalbandspektrogramm und Annotationstier mit Vorschlägen für laute (e) und lautlose (i) Anteile eines Säuglings mit sensorineuraler Hörschädigung vor Versorgung mit Hörhilfen, 79 Tage

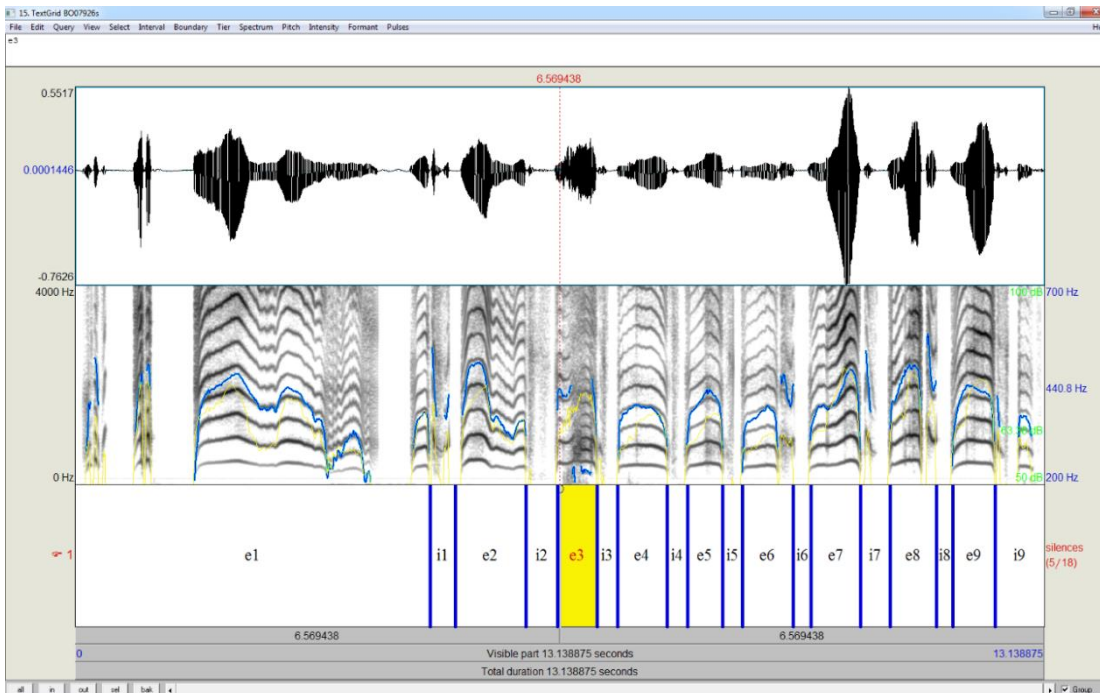


Abbildung 40: Testsignal H nach manueller Korrektur der Annotation von oben nach unten: Zeitbereichsdarstellung, Schmalbandspektrogramm und angepasste sowie nummerierte Annotationen für laute (e) und lautlose (i) Anteile eines Kindes mit sensorineuraler Hörschädigung vor Versorgung mit Hörhilfen, 79 Tage

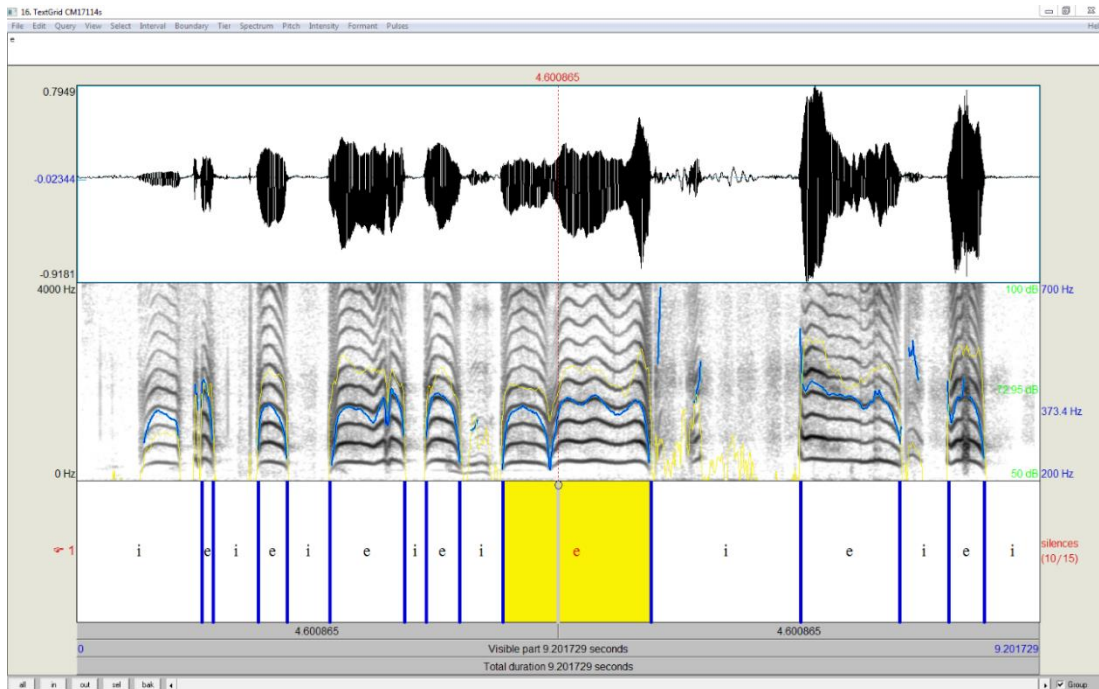


Abbildung 41: Testsignal I: automatisch generierter Annotationsvorschlag von oben nach unten: Zeitbereichsdarstellung, Schmalbandspektrogramm und Annotationstier mit Vorschlägen für laute (e) und lautlose (i) Anteile eines Kindes mit sensorineuraler Hörschädigung nach Versorgung mit Hörhilfen 171 Tage

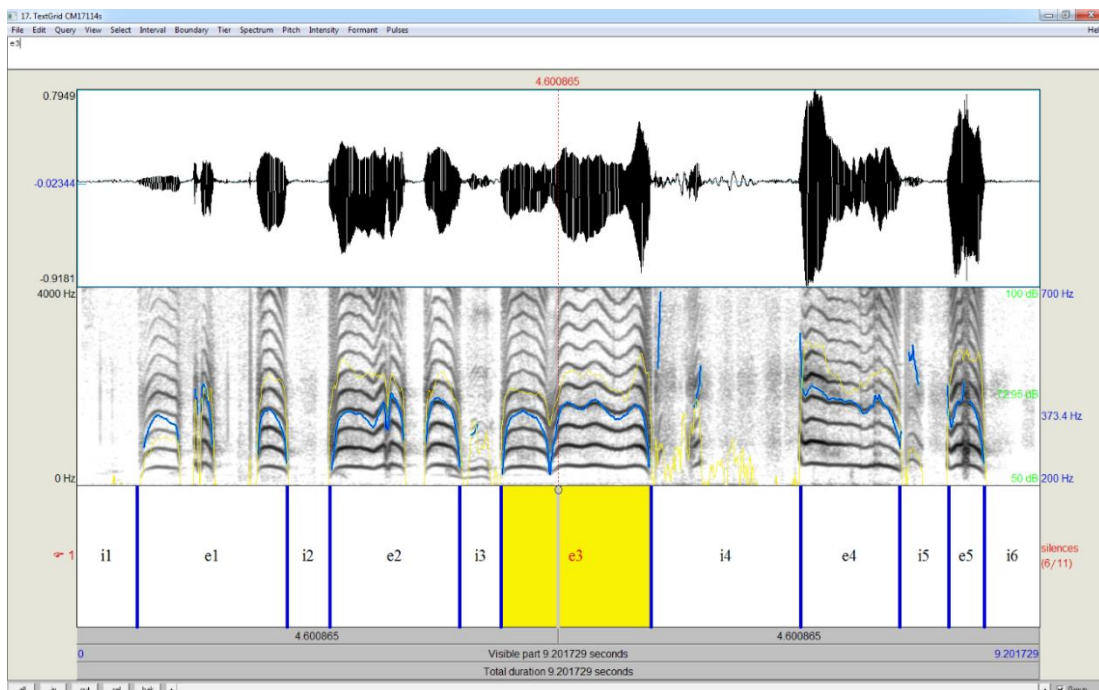


Abbildung 42: Testsignal I: manuell korrigierte Annotation von oben nach unten: Zeitbereichsdarstellung, Schmalbandspektrogramm und angepasste sowie nummerierte Annotationen für laute (e) und lautlose (i) Anteile eines Kindes mit sensorineuraler Hörschädigung nach Versorgung mit Hörhilfen, 171 Tage

#### 3.2.1.4 Ergebnisse der skriptbasierten Annotation der Testsignale aus der Entwicklungsphase des kanonischen Babbels unauffälliger Säuglinge

Dieser Annotationstest wurde an der in Kapitel 2.3.4 vorgestellten Lautsequenz des Testsignals J vorgenommen.

In diesem Beispiel wurde lediglich vor dem ersten Laut ein inkorrekt stiller Anteil entfernt. Die restlichen automatisch ermittelten und durch einen Cursor markierten Lautgrenzen waren korrekt und wurden in die finale Annotation übernommen, nummeriert und abgespeichert (Abbildung 43 vs. Abbildung 44).

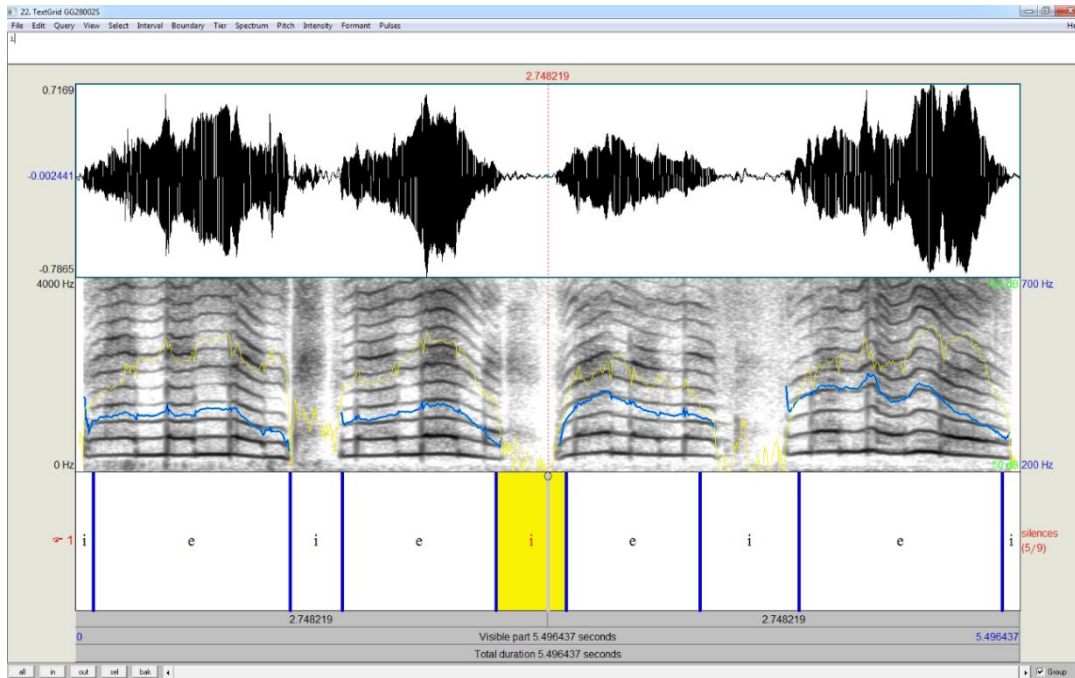


Abbildung 43: Testsignal J: automatisch generierter Annotationsvorschlag von oben nach unten: Zeitbereichsdarstellung, Schmalbandspektrogramm und Annotationstier mit Vorschlägen für laute (e) und lautlose (i) Anteile eines Kindes in der Entwicklungsphase des kanonischen Babelns 280 Tage

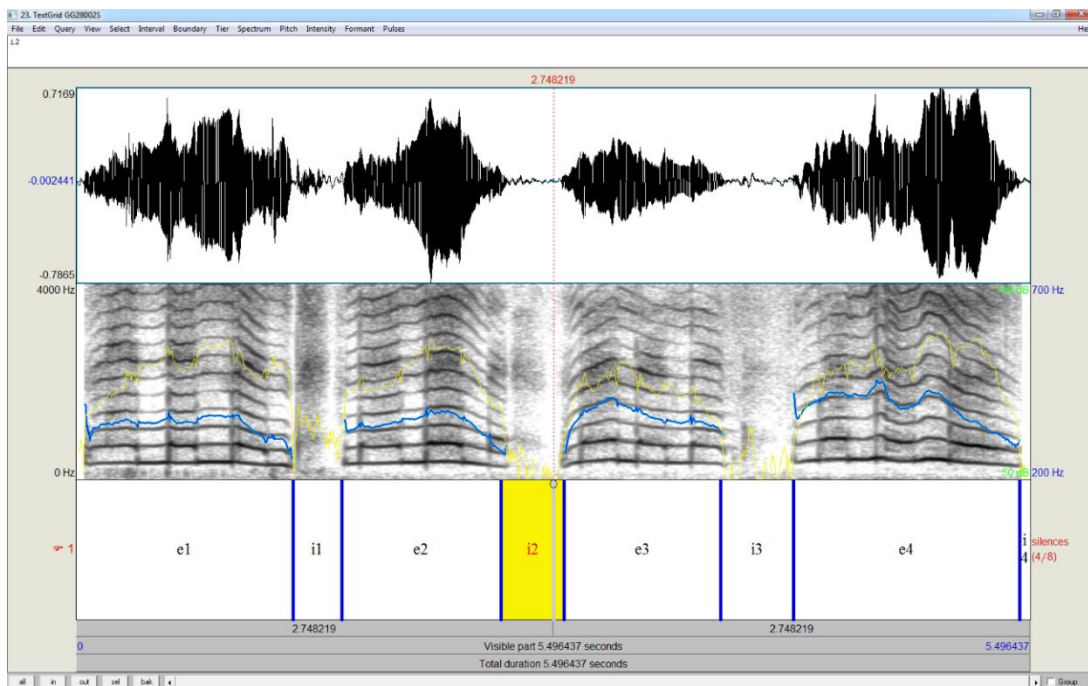


Abbildung 44: Testsignal J: vom Skript durchnummerierte Annotation von oben nach unten: Zeitbereichsdarstellung, Schmalbandspektrogramm und angepasste sowie nummerierte Annotationen für laute (e) und lautlose (i) Anteile eines Kindes in der Entwicklungsphase des kanonischen Babelns 280 Tage

### 3.2.2 Ermittlung temporaler Eigenschaften der Testsignale mithilfe des Skriptes zur Analyse annotierter Langserien

Ziel des hier erstellten Skriptes war auch eine Zeitersparnis bei der Vorverarbeitung von Signalen, die dann im zentrumseigenen Analysesystem für die Vorsprachliche Diagnostik des ZVES (CDAP-System, Autor: P. Wermke) weiterverarbeitet werden. Im CDAP-System werden unter anderem Kenngrößen der Laute analysiert, die Eigenschaften der Grundfrequenz ( $f_0$ ) erfassen, sowie ebenfalls temporale Größen, wie z.B. die Einzellaut- oder Segmentierungspausenlänge.

Auch die Analysen unter temporalen Gesichtspunkten werden bisher häufig unter Verwendung der  $f_0$ -Kurven durchgeführt. In der „klassischen“ Schreiforschung erfolgte die Analyse temporaler Lauteigenschaften sogar am Frequenzspektrum (z.B. Wasz-Hockert et al., 1968; Wasz-Höckert et al., 1985). Das hier erarbeitete Skript zur Analyse der annotierten Langserien bietet die Möglichkeit, direkt vom Oszillogramm (Zeitbereichsdarstellung) die Lautlängen aller annotierten Abschnitte abzugreifen (vgl. Kapitel 3.1.1.7 und Codeabschnitt 8).

Von allen in Kapitel 2.3 beschriebenen Testsignalen wurden die Lautlängen der Sequenzen jeweils am Zeitsignal mithilfe des Skriptes durch semi-automatische Cursorsetzung ermittelt und vergleichend an der Melodie im CDAP-System (bisheriges Vorgehen) bestimmt. Zwischen den in PRAAT per Skript am Zeitsignal ermittelten Werten und den an der Melodiekurve ermittelten Werten im CDAP-System wurden Differenzen gebildet. In Tabelle 2 sind die Ergebnisse der deskriptiven Statistik für den Vergleich dieser Zeitwerte aufgeführt. Im arithmetischen Mittel betrug die Differenz zwischen den am Zeitsignal bzw. der Melodie ermittelten Lautlängen 24,75ms ( $\pm 38$ ms), im Median 21ms.

Tabelle 2: Deskriptive Statistik der absoluten Zeitdifferenz der Lautlängen (Lges [s]) bei Messung am Zeitsignal bzw. der Melodiekontur

		Statistik	Std.-Fehler
Differenz	Mittelwert	24,75	5,504
Melodie, Zeitdarstellung	95% Konfidenzintervall des Mittelwerts	Untergrenze	13,67
		Obergrenze	35,83
	5% getrimmtes Mittel	25,99	
	Median	21,00	
	Std.-Abweichung	37,735	
	Minimum	-92	
	Maximum	97	

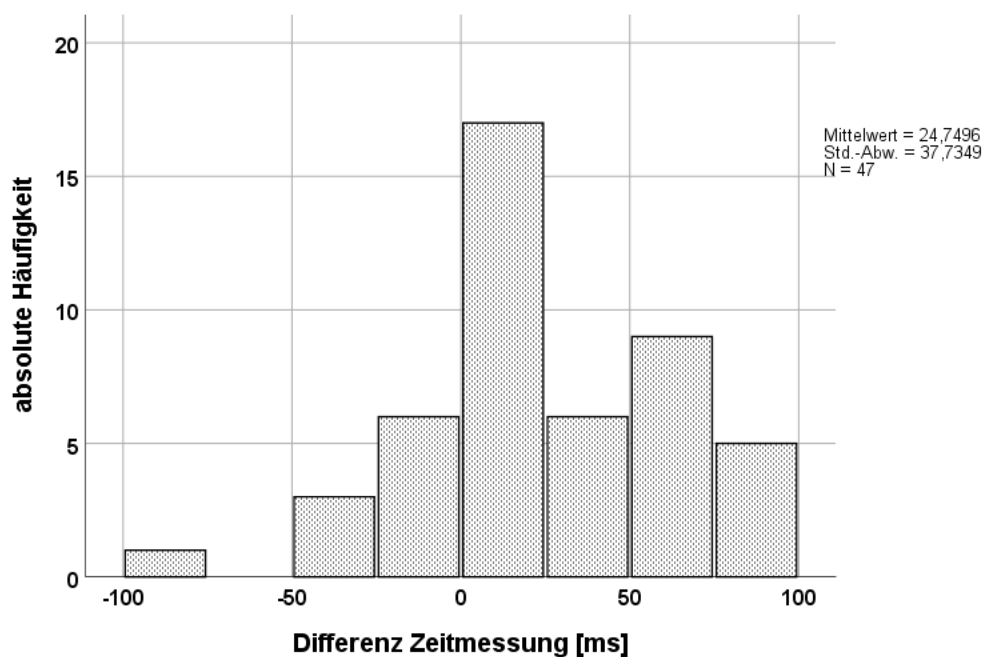


Abbildung 45: Histogramm der Verteilung der Messdifferenzen für die Testsignale im Methodenvergleich

Wie in den Verteilungseigenschaften in Abbildung 45 zu erkennen ist, befindet sich der überwiegende Teil der Werte im positiven Bereich. Auch der Median liegt mit 21ms im positiven Bereich. Das bedeutet, dass die Werte für die

Vokalisationslängen, die direkt am Zeitsignal ermittelt wurden, im Durchschnitt geringfügig länger waren. Ergänzend zu diesem Vergleich wurden für alle Zeitmessungen im Rahmen der Projektauswertungen (Kap. 3.3) sowohl die bisherige Methode als auch die hier neu erarbeitete und getestete verwendet. Dadurch liegen zusätzliche Methodenvergleiche vor.

### 3.3 Teil 3: Anwendung des erarbeiteten Skriptes auf Projektdaten

#### 3.3.1 Vergleich der Melodiestrukturkategorien in Lauten von Säuglingen mit und ohne orofaziale Spaltbildung

Das vom Autor erarbeitete und getestete Hauptskript kann zur Analyse temporaler Größen verschiedener Strukturen und Elemente eingesetzt werden. Neben den Gesamtlängen von Äußerungen können auch kürzere zeitliche Einheiten wie Melodie- oder Intonationsbögen, Silbenlängen, innerlautliche Pausen oder Pausen zwischen einzelnen Vokalisationen, Phrasen oder Sätzen ermittelt werden. In dem Projekt (Kapitel 2.1) dessen Daten hier analysiert wurden, sollten temporale Größen in Melodiestrukturen im spontanen Säuglingsweinen analysiert werden. Dazu mussten die Melodiestrukturen zunächst mit Hilfe der am ZVES verwendeten Routineverfahren (CDAP) erfasst werden. Dabei wird zwischen einbögigen (1B) und mehrbögigen (2B, 3B, MB) bzw. mehrbödig-segmentierten, d.h. durch innerlautliche Pausen charakterisierte, Strukturen (1S, KS, 2S, 3S, MS) unterschieden (fuzzy class „?“). Die Ergebnisse sind in Abbildung 46 berichtet.

In Abbildung 46 sind die Auftrittshäufigkeiten der Melodiestrukturkategorien der Laute beider Probandengruppen des Projektes in Form eines Säulendiagramms dargestellt.



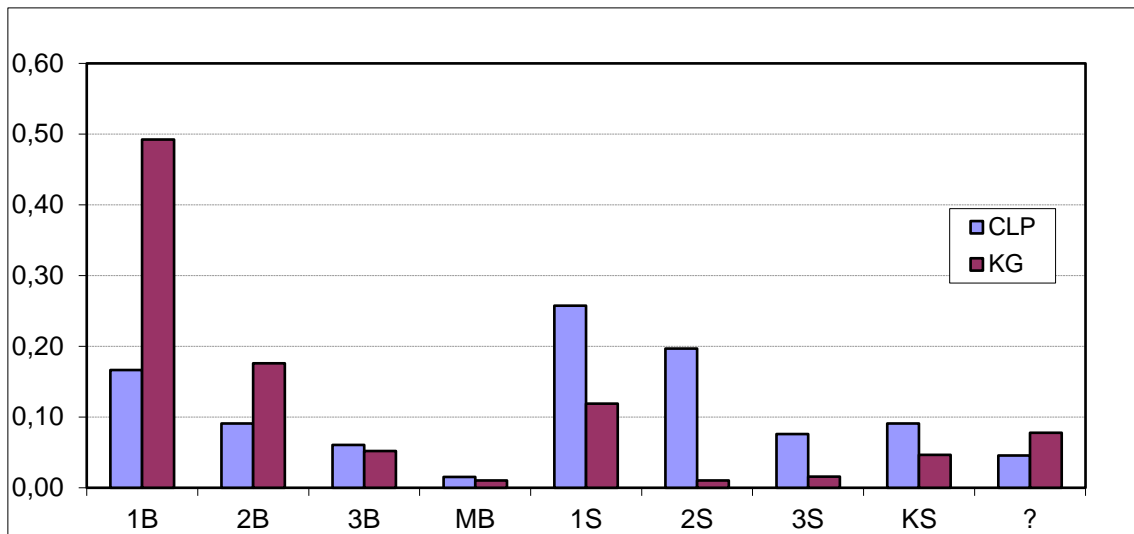


Abbildung 46: Säulendiagramm zum Vergleich der relativen Auftrittshäufigkeiten der Melodiestrukturkategorien in Säuglingslauten der Probanden mit (CLP, blau) und ohne (KG, lila) orofaziale Spaltbildungen

Die zugehörigen (absoluten und relativen) Werte pro Strukturkategorie und Gruppe sind in Tabelle 3 berichtet.

Tabelle 3: Absolute (oben) und relative (unten) Häufigkeiten der Melodiestrukturkategorien in beiden Probandengruppen

Gruppe	ges N	1B	2B	3B	MB	1S	2S	3S	KS	?
CLP	66	11	6	4	1	17	13	5	6	3
KG	193	95	34	10	2	23	2	3	9	15
CLP		0,17	0,09	0,06	0,02	0,26	0,20	0,08	0,09	0,05
KG		0,49	0,18	0,05	0,01	0,12	0,01	0,02	0,05	0,08

Die Beobachtungen belegen, dass der Anteil innermelodisch segmentierter Kategorien (\*S) in den Vokalisationen der Gruppe der LKGS-Probanden (CLP) im Vergleich zur gleichaltrigen Kontrollgruppe (63% vs 20%) deutlich höher ist.

Die pro Säugling berechneten und dann über die Gruppe gemittelten Indices (Kapitel 2.1) ergaben folgende Werte:

Der Melodiekomplexitätsindex (MCI) wurde mit 0,47 für die KG-Gruppe und 0,83 für die LKGS-Gruppe ermittelt. Der Segmentierungsindex (SI) beträgt 0,45 für die KG-Gruppe und 0,79 für die LKGS-Gruppe.

### 3.3.2 Vergleich der temporalen Eigenschaften in Säuglingslauten der Probandengruppen des Projektes

In Abbildung 47 sind die Verteilungseigenschaften der Gesamtlängen aller analysierten Vokalisationen für beide Gruppen im Boxplot dargestellt (SAM-Methode). In Tabelle 4 ist die zugehörige deskriptive Statistik berichtet.

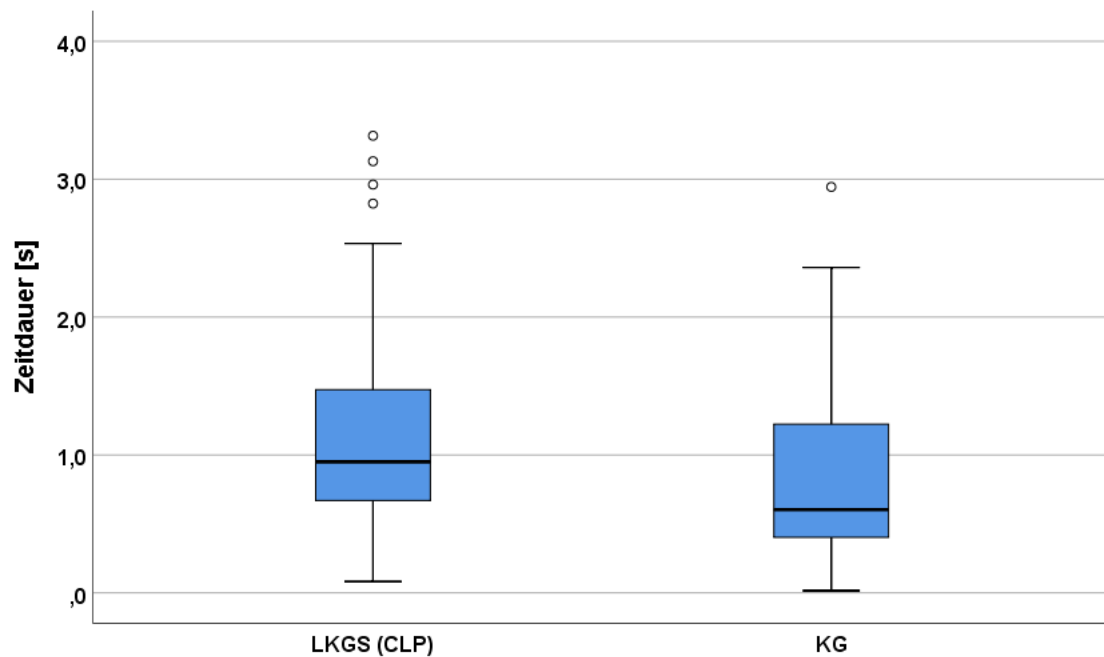


Abbildung 47: Boxplotdiagramm der Verteilungseigenschaften der Gesamtlängen (Lges [s]) aller analysierten Vokalisationen (SAM-Methode)

Tabelle 4: Deskriptive Statistik der Verteilungseigenschaften der Gesamtlängen (Lges [s]) aller analysierten Vokalisationen (SAM-Methode)

Gruppe		Statistik	Std.-Fehler	
CLP	Mittelwert	1,151124	,0957919	
	95% Konfidenzintervall des Mittelwerts	Untergrenze	,959814	
		Obergrenze	1,342433	
	5% getrimmtes Mittel	1,098518		
	Median	,950250		
	Varianz	,606		
	Std.-Abweichung	,7782169		
	Minimum	,0819		
	Maximum	3,3148		
	KG	Mittelwert	,816390	,0434130
95% Konfidenzintervall des Mittelwerts		Untergrenze	,730760	
		Obergrenze	,902021	
5% getrimmtes Mittel		,780646		
Median		,603532		
Varianz		,362		
Std.-Abweichung		,6015478		
Minimum		,0161		
Maximum		2,9437		
Spannweite		2,9277		

In den nachfolgenden Unterkapiteln werden die in Kapitel 2.1 definierten temporalen Eigenschaften für jede der Melodiestrukturkategorien in den Säuglingslauten beider Projektgruppen berichtet (Signallevel). Zusätzlich sind auch hier noch einmal vergleichend die Ergebnisse anhand der Messung an der Melodiekontur (MM-Methode), also dem bisherigen Routineverfahren am ZVES vor der Skriptanwendung, mitberichtet. In den entsprechenden MM-Tabellen gibt es für einzelne Strukturen, im Unterschied zu den SAM-Tabellen, Fehlwerte. Der Grund dafür ist, dass die Zeitmessung an der Melodie nur funktioniert, wenn die Melodie vollständig berechnet vorlag; in stimmlosen, verrauschten oder anderweitig gestörten Signalbereichen kommen häufig Grundfrequenzfehlwerte vor.

### 3.3.2.1 Temporale Eigenschaften einböiger (1B) Strukturen

In Tabelle 7 sind die Ergebnisse der Messungen für alle Laute mit einböiger Melodie (1B) auf Signallevel pro Gruppe und Säugling (sofern der jeweilige Säugling Laute dieser Strukturkategorie erzeugt hat) berichtet. Dies war im Projekt gefordert, um erste Referenzwerte für nachfolgende Studien zu dokumentieren. Es sind sowohl die am Zeitsignal mit der hier entwickelten semi-automatischen Methode (SAM) ermittelten Gesamtlängen als auch die auf Vermessung der Melodiekurven basierenden Lautlängen (MM) berichtet. Da in der semi-automatischen Methode die Kontextinformation des jeweiligen Signals erhalten bleibt, konnten auch das jeweilig vorangehende und nachfolgende Inspirationssignal mit vermessen und berichtet werden. Bei der melodiebasierten Methode war dies nicht möglich, da nur Einzelvokalisationen in den Audiofiles gespeichert vorliegen.

In Abbildung 48 sind die Verteilungseigenschaften der Werte für die Vokalisationslängen (1B) noch einmal grafisch in einem Boxplotdiagramm dargestellt (Signallevel). Die zugehörige deskriptive Statistik ist in Tabelle 5 für die SAM Methode und in Tabelle 6 für die MM Methode berichtet.

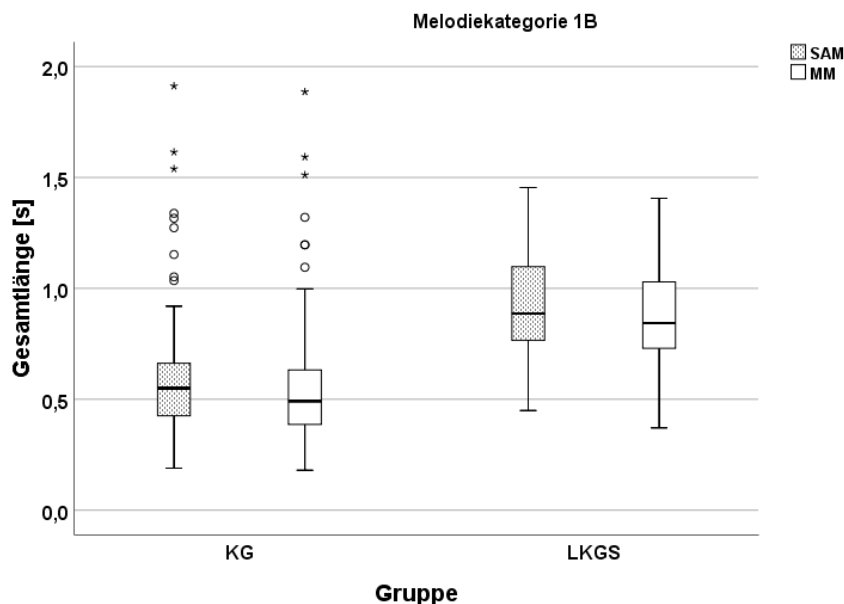


Abbildung 48: Boxplotdiagramm: Vergleich der Gesamtlängen (Lges [s]) von 1B-Strukturen im Gruppen- und Methodenvergleich

Tabelle 5: Deskriptive Statistik für die Gesamtlängen (Lges [s]); SAM-Methode im Gruppenvergleich für 1B- Strukturen

	Gruppe		Statistik	Std.-Fehler	
SAM	KG	Mittelwert	,60513	,030295	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	,54498	
			Obergrenze	,66528	
		5% getrimmtes Mittel	,57048		
		Median	,55037		
		Varianz	,087		
		Std.-Abweichung	,295276		
		Minimum	,190		
		Maximum	1,912		
	LKGS	Mittelwert	,93776	,084116	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	,75034	
			Obergrenze	1,12518	
		5% getrimmtes Mittel	,93618		
		Median	,88627		
		Varianz	,078		
		Std.-Abweichung	,278982		
		Minimum	,449		
		Maximum	1,455		

Tabelle 6: Deskriptive Statistik für die Gesamtlängen (Lges [s]); MM-Methode im Gruppenvergleich für 1B-Strukturen

	Gruppe		Statistik	Std.-Fehler	
MM	KG	Mittelwert	,56717	,030017	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	,50757	
			Obergrenze	,62677	
		5% getrimmtes Mittel	,53180		
		Median	,49125		
		Varianz	,086		
		Std.-Abweichung	,292568		
		Minimum	,180		
		Maximum	1,886		
	LKGS	Mittelwert	,87941	,083782	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	,69273	
			Obergrenze	1,06609	
		5% getrimmtes Mittel	,87837		
		Median	,84375		
		Varianz	,077		
		Std.-Abweichung	,277875		
		Minimum	,371		
		Maximum	1,406		

Tabelle 7: Messwerte der Vokalisationslängen (Lges [s]) mit der semi-automatischen Methode (SAM) und der manuellen Methode (MM) für Strukturen des Typs 1B (I1 und I2 = Inspirationslängen [s]; Pb = Proband)

Gruppe	Pb	Signal	I1	SAM	I2	MM
LKGS	1	AA08205m	0,216	0,449	0,435	0,371
LKGS	1	AA08206m	0,528	1,094	0,223	1,016
LKGS	1	AA08224c	0,500	0,851	0,256	0,784
LKGS	1	AA08226c	0,437	0,886	0,236	0,844
LKGS	3	AC05208c	0,393	1,455	0,238	1,406
LKGS	3	AC05209c	0,269	1,241	0,260	1,181
LKGS	3	AC05210c	0,232	1,102	0,227	1,043
LKGS	3	AC05213c	0,287	1,001	0,301	0,900
LKGS	3	AC05214c	0,292	0,788	0,192	0,728
LKGS	3	AC05224c	0,317	0,703	0,286	0,669
LKGS	4	AD06213c	0,300	0,745	0,246	0,732
KG	5	AA05631c	0,549	0,805	1,464	0,630
KG	5	AA05638c	0,741	0,705	0,552	0,593
KG	5	AA05639c	0,731	0,835	1,171	0,806
KG	5	AA05641c	0,307	1,152	1,298	1,095
KG	5	AA05679m	0,223	0,489	0,287	0,454
KG	5	AA05686c	0,249	1,035		0,990
KG	5	AA05678n	0,152	0,675	0,220	0,660
KG	6	AB05809m	0,225	0,336	0,234	0,330
KG	6	AB05817m	0,639	0,550	0,505	0,525
KG	6	AB05818m	1,019	0,526	0,597	0,458
KG	6	AB05821m	0,299	0,190	0,268	0,368
KG	6	AB05823m	0,053	0,438	0,223	0,308
KG	6	AB05825c	0,082	0,570	0,329	0,443
KG	6	AB05828m	0,205	0,893	0,271	0,784
KG	6	AB05836c	0,403	0,557	0,105	0,529
KG	6	AB05840m	0,433	0,761	0,302	0,750
KG	6	AB05841m	0,184	0,558	0,236	0,536
KG	6	AB05845c	0,118	0,485	0,277	0,383
KG	6	AB05847c	0,477	0,566		0,473
KG	6	AB05850m		0,678		0,656
KG	6	AB05853m	0,195	0,515	0,208	0,495
KG	6	AB05854m	0,194	0,369	0,195	0,356
KG	6	AB05855m	0,177	0,362	0,163	0,345
KG	6	AB05856m	0,138	0,384	0,237	0,368
KG	6	AB05867n	0,106	0,260	0,189	0,180
KG	6	AB05880m	0,166	0,425	0,146	0,375
KG	6	AB05886m	0,177	0,625	0,210	0,525
KG	6	AB058a0c	0,111	0,553	0,302	0,401
KG	6	AB058a4m	0,179	0,467	0,253	0,428
KG	6	AB058a5m	0,141	0,446	0,203	0,409
KG	6	AB058a8m	0,130	0,310	0,225	0,334
KG	6	AB058b1m	0,226	0,596	0,197	0,540
KG	6	AB058b3c	0,075	0,404	0,346	0,338
KG	6	AB058b4c	0,082	0,408	0,392	0,401
KG	6	AB05830c	0,235	0,545	0,351	0,405
KG	6	AB05842m	0,178	0,699	0,120	0,686
KG	6	AB05844m	0,171	0,572	0,143	0,559

Gruppe	Pb	Signal	I1	SAM	I2	MM
KG	7	AC05601m	0,183	0,428		0,278
KG	7	AC05602m	0,263	0,439		0,405
KG	7	AC05603m	0,289	0,447		0,405
KG	7	AC05604m	0,227	1,539		1,511
KG	7	AC05610c	0,029	0,488	0,256	0,480
KG	7	AC05611c	0,042	0,504	0,222	0,484
KG	7	AC05612c	0,043	0,558	0,271	0,548
KG	7	AC05613c	0,063	0,663	0,224	0,645
KG	7	AC05614c	0,054	0,590	0,224	0,555
KG	7	AC05615c	0,052	0,571	0,250	0,559
KG	7	AC05616m	0,037	0,602	0,240	0,593
KG	7	AC05617m	0,116	0,608	0,227	0,585
KG	7	AC05618m	0,032	0,688	0,240	0,679
KG	7	AC05619m	0,068	0,661	0,221	0,649
KG	7	AC05620m	0,089	0,643	0,204	0,634
KG	7	AC05621m	0,025	0,637	0,221	0,626
KG	7	AC05622m	0,074	0,895	0,220	0,881
KG	7	AC05623m	0,061	0,700	0,282	0,694
KG	7	AC05625m	0,039	0,591	0,220	0,585
KG	7	AC05626m	0,081	0,412	0,295	0,416
KG	7	AC05627m	0,031	0,806	0,267	0,795
KG	7	AC05629m	0,030	0,363	0,177	0,356
KG	7	AC05630m	0,118	0,632	0,476	0,596
KG	7	AC05633m	0,059	0,562	0,536	0,521
KG	7	AC05634m	0,031	0,919	0,232	0,904
KG	8	AD06902m	0,292	1,273		1,196
KG	8	AD06905m	0,428	0,615	0,365	0,563
KG	8	AD06907m	0,418	0,403		0,338
KG	8	AD06908m	0,413	0,400		0,398
KG	8	AD06913c	0,322	1,912	0,290	1,886
KG	8	AD06916c	0,173	1,316	0,257	1,196
KG	8	AD06917c	0,247	1,338	0,207	1,320
KG	8	AD06920c	0,167	1,614	0,239	1,593
KG	8	AD06927m	0,288	0,452	0,291	0,443
KG	8	AD06928m	0,258	0,402	0,252	0,356
KG	8	AD06929m	0,283	0,448	0,315	0,431
KG	8	AD06930m	0,268	0,377	0,266	0,345
KG	8	AD06937m	0,270	0,464	0,284	0,423
KG	8	AD06938m	0,145	0,320	0,170	0,268
KG	8	AD06943m	0,790	0,474	0,725	0,428
KG	8	AD06946m	0,334	0,387	0,256	0,371
KG	8	AD06953m	0,295	0,465	0,213	0,461
KG	8	AD06954m	0,229	0,519	0,225	0,503
KG	8	AD06963m	0,218	0,582	0,327	0,491
KG	8	AD06964m	0,353	0,523	0,225	0,521
KG	8	AD06965m	0,227	0,427	0,192	0,420
KG	8	AD06967m	0,264	0,504	0,206	0,484
KG	8	AD06971m	0,215	0,322	0,206	0,255
KG	8	AD06974m	0,348	0,349	0,262	0,360
KG	8	AD06975m	0,163	0,578	0,377	0,559
KG	8	AD06982m	0,272	0,423	0,265	0,330
KG	8	AD06986m	0,084	0,403	0,163	0,355
KG	8	AD06994m	0,168	0,353	0,193	0,353

Gruppe	Pb	Signal	I1	SAM	I2	MM
KG	8	AD069a3m	0,264	0,430	0,249	0,390
KG	8	AD069a4m	0,216	0,605	0,309	0,548
KG	8	AD069a6c	0,205	1,052	0,249	0,998
KG	8	AD069c0m	0,132	0,326	0,115	0,300
KG	8	AD069d3m	0,473	0,738	0,106	0,709

Zwischen den Methoden zeigte sich für Lges eine Differenz der Mittelwerte von 0,038s in der Kontrollgruppe (Mediandifferenz 0,059s). In der LKGS-Gruppe betrug die Differenz der Mittelwerte von Lges auf Signallevel 0,058s (Mediandifferenz 0,043s).

Die deskriptive Statistik der absoluten Zeitmessungsdifferenz auf Signallevel zwischen den Vokalisationen vom **Typ 1B** ist in der Tabelle 8 berichtet (Einzelwerte im Anhang A 2). Abbildung 49 zeigt die Verteilung der absoluten Zeitmessungsdifferenz zwischen beiden Methoden für beide Probandengruppen in Form eines Boxplotdiagrammes.

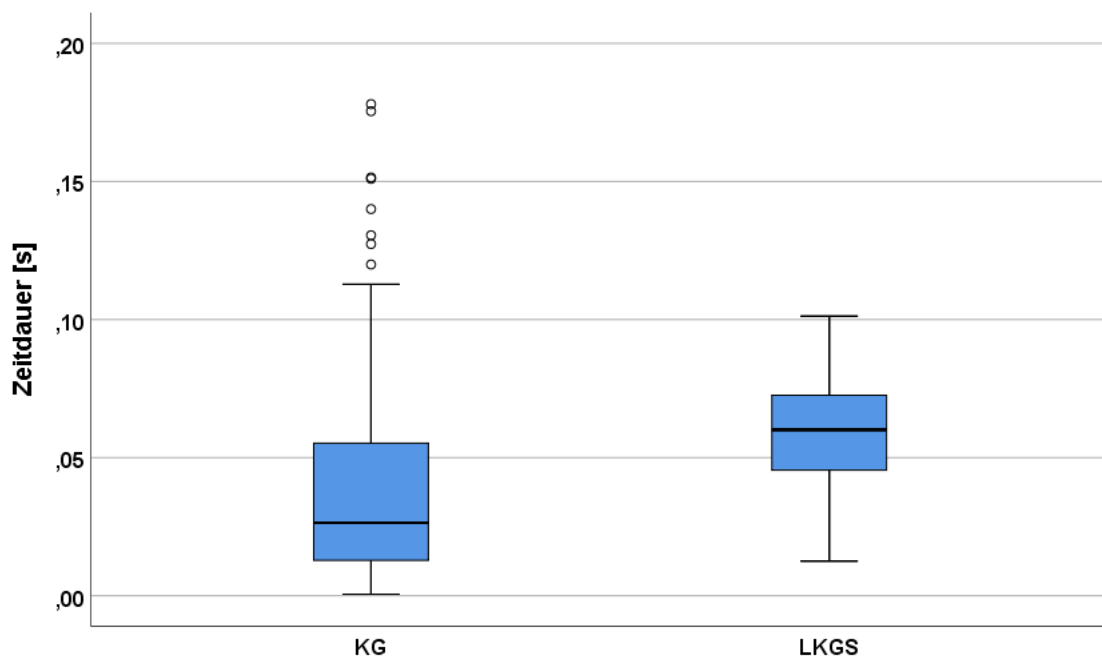


Abbildung 49: Boxplotdiagramm der absoluten Zeitmessungsdifferenz zwischen SAM- und MM-Methode für beide Gruppen (Vokalisationstyp 1B)



Tabelle 8: Deskriptive Statistik für die absolute Zeitmessungsdifferenz [s] der Gesamtlänge der Vokalisationen des Typs 1B zwischen SAM- und MM-Methode beider Gruppen

	Gruppe		Statistik	Std.-Fehler	
ABSDiff_L1B	KG	Mittelwert	,04252	,004283	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	,03402	
			Obergrenze	,05103	
		5% getrimmtes Mittel	,03821		
		Median	,02643		
		Varianz	,002		
		Std.-Abweichung	,041749		
		Minimum	,001		
		Maximum	,178		
	LKGS	Mittelwert	,05835	,007225	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	,04225	
			Obergrenze	,07445	
		5% getrimmtes Mittel	,05852		
		Median	,06010		
		Varianz	,001		
		Std.-Abweichung	,023964		
		Minimum	,013		
		Maximum	,101		

### 3.3.2.2 Temporale Eigenschaften nichtsegmentierter zweibögiger Strukturen (2B)

Abbildung 50 zeigt die Verteilung der Gesamtlängen im Gruppen- und Methodenvergleich für die doppelbögigen Melodiestruckturen (Signallevel) in Form eines Boxplotdiagramms. Die zugehörige deskriptive Statistik ist in Tabelle 9 für die SAM-Methode und in Tabelle 10 für die MM-Methode berichtet. In Tabelle 11 sind alle Messwerte der Vokalisationen (Säuglingsschreie) mit zweibögiger Melodiekontur pro Säugling (sofern produziert) und Gruppe berichtet. Für die SAM-Messungen sind auch hier wieder die der jeweiligen Vokalisation vorangehenden bzw. nachfolgenden Inspirationslängen (soweit eindeutig vermessenbar) mit aufgeführt.

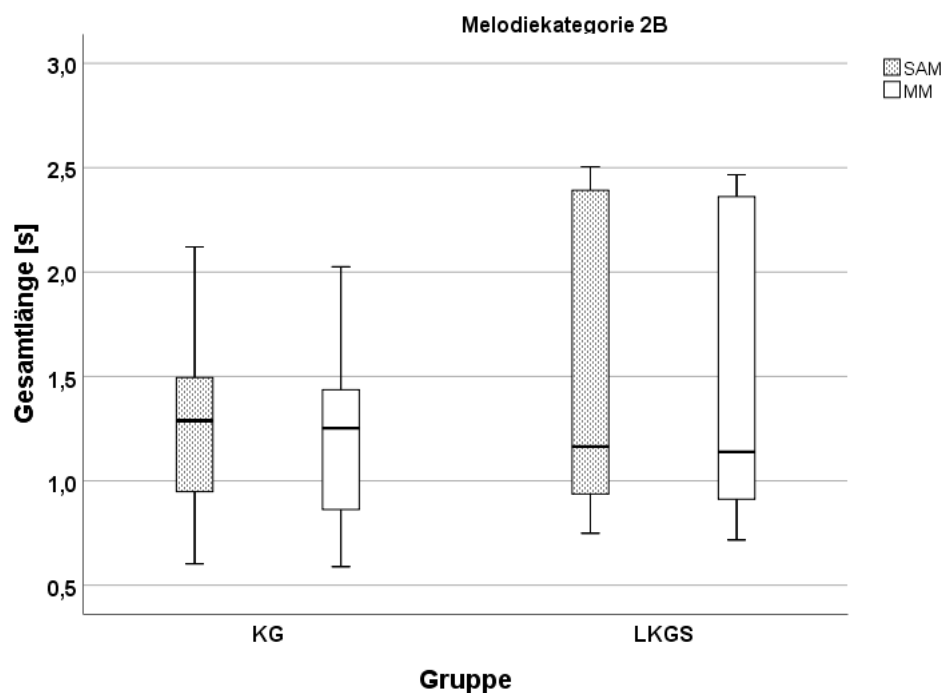


Abbildung 50: Boxplotdiagramm: Vergleich der Gesamtlängen (Lges [s]) von 2B-Strukturen im Gruppen- und Methodenvergleich

Tabelle 9: Deskriptive Statistik für die Gesamtlängen (Lges [s]); SAM-Methode im Gruppenvergleich für 2B-Strukturen

Gruppe		Statistik	Std.-Fehler		
SAM	KG	Mittelwert	1,29750	,067271	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	1,16064	
			Obergrenze	1,43437	
		5% getrimmtes Mittel	1,28856		
		Median	1,28835		
		Varianz	,154		
		Std.-Abweichung	,392256		
		Minimum	,602		
		Maximum	2,120		
	LKGS	Mittelwert	1,48494	,315238	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	,67459	
			Obergrenze	2,29528	
		5% getrimmtes Mittel	1,46917		
		Median	1,16385		
		Varianz	,596		
		Std.-Abweichung	,772171		
Minimum	,749				
Maximum	2,505				

Tabelle 10: Deskriptive Statistik für die Gesamtlängen (Lges [s]); MM-Methode im Gruppenvergleich für 2B-Strukturen

	Gruppe		Statistik	Std.-Fehler	
MM	KG	Mittelwert	1,23073	,065730	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	1,09700	
			Obergrenze	1,36446	
		5% getrimmtes Mittel	1,22142		
		Median	1,25250		
		Varianz	,147		
		Std.-Abweichung	,383270		
		Minimum	,589		
		Maximum	2,025		
	LKGS	Mittelwert	1,45557	,313643	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	,64932	
			Obergrenze	2,26182	
		5% getrimmtes Mittel	1,44048		
		Median	1,13888		
		Varianz	,590		
		Std.-Abweichung	,768266		
		Minimum	,717		
		Maximum	2,466		

Tabelle 11: Messwerte aller Vokalisationslängen (Lges [s]) mit der semi-automatischen Methode (SAM) und der manuellen Methode (MM) für Strukturen des Typs 2B (I1 und I2 [s] = Inspirationslängen)

Gruppe	Pb	Signal	I1	SAM	I2	MM
LKGS	3	AC05201c	0,260	1,353	0,227	1,324
LKGS	3	AC05211c	0,265	0,749	0,202	0,717
LKGS	3	AC05221c	0,375	0,937	0,285	0,912
LKGS	3	AC05222c	0,309	0,974	0,277	0,954
LKGS	4	AD06206c	0,231	2,505	0,251	2,466
LKGS	4	AD06207c	0,308	2,391	0,305	2,361
KG	5	AA05642c	0,373	0,770	1,388	0,713
KG	5	AA05647c	0,172	1,911	0,563	1,868
KG	5	AA05650c	0,429	1,963	0,837	1,939
KG	5	AA05651m	0,289	1,398	0,595	1,376
KG	5	AA05652c	0,361	1,305	0,982	1,260
KG	5	AA05656c	0,158	1,463	0,198	1,406
KG	5	AA05659c	0,221	1,398	0,587	1,313
KG	5	AA05660c	0,165	1,494	0,376	1,436
KG	5	AA05670c	0,114	1,284	0,374	1,260
KG	5	AA05675c	0,317	1,369	0,523	1,328
KG	5	AA05681c	0,276	1,301	0,248	1,268
KG	5	AA05682c	0,246	1,775	1,202	1,688
KG	5	AA05657u	0,191	1,401	0,227	1,297
KG	6	AB05815m	0,198	0,827	0,244	0,761

Gruppe	Pb	Signal	I1	SAM	I2	MM
KG	6	AB05837c	0,388	1,901		1,864
KG	6	AB05851m	0,171	1,182	0,311	1,168
KG	6	AB05852m	0,169	0,929	0,273	0,919
KG	6	AB05870m	0,264	0,906	0,290	0,848
KG	6	AB058b7c	0,155	1,536		1,481
KG	7	AC05632m	0,070	2,034	0,480	2,025
KG	7	AC05637m	0,043	0,840	0,256	0,821
KG	8	AD06903m	0,317	1,292		1,245
KG	8	AD06904m	0,361	0,899	0,272	0,863
KG	8	AD06914c	0,522	1,664	0,403	1,635
KG	8	AD06932m	0,663	0,948	0,943	0,806
KG	8	AD06949m	0,313	1,050	0,220	1,035
KG	8	AD06959m	0,236	1,091	0,696	1,024
KG	8	AD06984c	0,237	1,192	0,581	1,069
KG	8	AD06992m	0,078	0,930	0,183	0,746
KG	8	AD06998m	0,329	0,602	0,260	0,589
KG	8	AD06999c	0,185	1,238	0,422	1,226
KG	8	AD069b7c	0,299	2,120	0,469	1,676
KG	8	AD069c3m	0,338	1,025	0,411	0,841
KG	8	AD069d4c	0,158	1,077	0,597	1,054

Zwischen den Methoden zeigte sich eine Differenz der Mittelwerte von 0,067s in der Kontrollgruppe (Mediandifferenz 0,036s). In der LKGS-Gruppe betrug die Differenz auf Signalleve im Mittel 0,029s (Mediandifferenz 0,025s).

Die deskriptive Statistik der absoluten Zeitmessungsdifferenz zwischen den Vokalisationen vom **Typ 2B** ist in Tabelle 12 berichtet (Einzelwerte im Anhang A 2). Hier ist nicht zwischen den Gruppen unterschieden worden, da es in der LKGS-Gruppe nur sechs Vokalisationen gab.

Tabelle 12: Deskriptive Statistik der absoluten Zeitmessungsdifferenzen [s] für die Gesamtlängen (Lges [s]) der Vokalisationen des Typs 2B zwischen SAM- und MM-Methode beider Gruppen

		Statistik	Std.-Fehler	
Diffabs	Mittelwert	,0612	,01195	
	95% Konfidenzintervall des Mittelwerts	Untergrenze	,0370	
		Obergrenze	,0853	
	5% getrimmtes Mittel	,0500		
	Median	,0379		
	Varianz	,006		
	Std.-Abweichung	,07557		
	Minimum	,01		
	Maximum	,44		

### 3.3.2.3 Temporale Eigenschaften drei- (3B) und mehrböiger (MB) Strukturen

Die drei- und mehrböigen Melodiekonturen sind vergleichsweise selten beobachtet worden und daher hier zusammengefasst berichtet. Die Messwerte (Signallevel) sind grafisch im Boxplot (Abbildung 51) und deskriptiv in Tabelle 13 (SAM) und Tabelle 14 (MM) berichtet. Die einzelnen Messwerte auf Signalebene zeigt Tabelle 15.

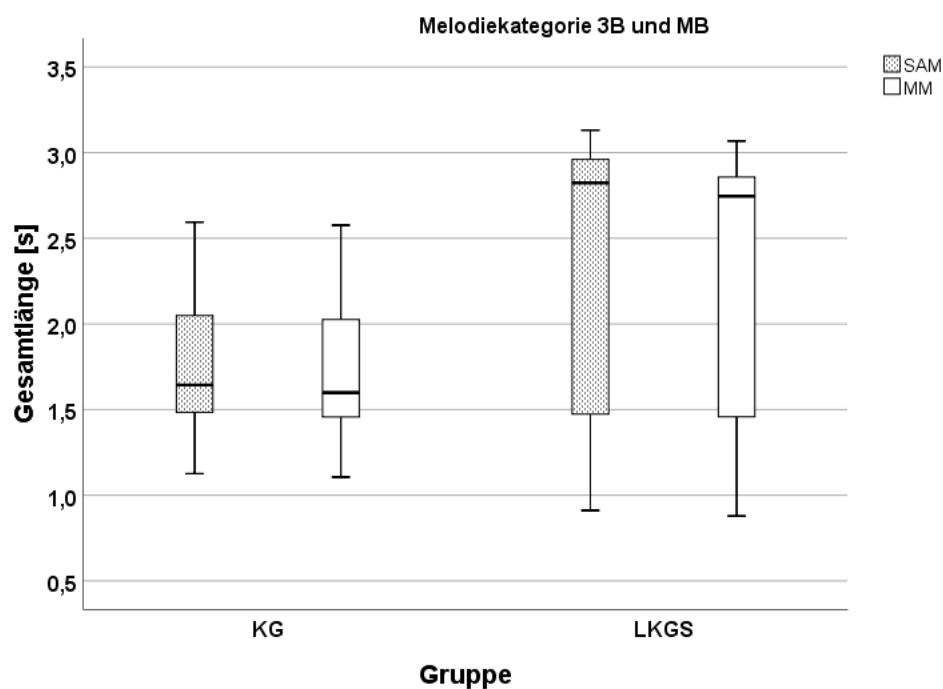


Abbildung 51: Boxplotdiagramm: Vergleich der Gesamtlängen (Lges [s]) von 3B- und MB- Strukturen im Gruppen- und Methodenvergleich

Tabelle 13: Deskriptive Statistik für die Gesamtlängen (Lges [s]); SAM-Methode im Gruppenvergleich für 3B- und MB-Strukturen

	Gruppe		Statistik	Std.-Fehler	
SAM	KG	Mittelwert	1,74499	,124348	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	1,47131	
			Obergrenze	2,01868	
		5% getrimmtes Mittel	1,73221		
		Median	1,64406		
		Varianz	,186		
		Std.-Abweichung	,430756		
		Minimum	1,127		
		Maximum	2,593		
	LKGS	Mittelwert	2,25976	,447459	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	1,01741	
			Obergrenze	3,50210	
		5% getrimmtes Mittel	2,28631		
		Median	2,82327		
		Varianz	1,001		
		Std.-Abweichung	1,000549		
		Minimum	,911		
		Maximum	3,131		

Tabelle 14: Deskriptive Statistik für die Gesamtlängen (Lges [s]); MM-Methode im Gruppenvergleich für 3B- und MB-Strukturen

	Gruppe		Statistik	Std.-Fehler	
MM	KG	Mittelwert	1,71719	,129984	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	1,43109	
			Obergrenze	2,00328	
		5% getrimmtes Mittel	1,70340		
		Median	1,59937		
		Varianz	,203		
		Std.-Abweichung	,450277		
		Minimum	1,106		
		Maximum	2,576		
	LKGS	Mittelwert	2,20140	,434597	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	,99477	
			Obergrenze	3,40803	
		5% getrimmtes Mittel	2,22675		
		Median	2,74500		
		Varianz	,944		
		Std.-Abweichung	,971787		
		Minimum	,879		
		Maximum	3,068		

Tabelle 15: Messwerte der Vokalisationslängen (Lges [s]) mit der semi-automatischen Methode (SAM) und der manuellen Methode (MM) für Strukturen des Typs 3B und MB (I1 und I2 [s] = Inspirationslängen, QSS = Strukturkategorie)

Gruppe	Pb	Signal	QSS	I1	SAM	I2	MM
LKGS	3	AC05215c	3B	0,186	0,911	0,288	0,879
LKGS	4	AD06202c	3B	0,189	2,823	0,308	2,745
LKGS	4	AD06204c	3B	0,322	3,131	1,514	3,068
LKGS	4	AD06212c	3B	0,283	1,473	0,275	1,458
KG	5	AA05629c	3B	0,619	1,127	1,775	1,106
KG	5	AA05653c	3B	0,267	1,494	0,594	1,451
KG	5	AA05674c	3B	0,320	1,645	0,654	1,616
KG	5	AA05677c	3B	0,278	1,743	0,552	1,706
KG	6	AB05888m	3B	0,263	1,474	0,259	1,463
KG	8	AD06918c	3B	0,346	2,001	0,147	1,990
KG	8	AD06922c	3B	0,346	1,644	0,279	1,553
KG	8	AD069b5c	3B	0,795	1,627	0,616	1,583
KG	8	AD069c7m	3B	0,266	1,209	0,585	1,145
KG	8	AD069d0c	3B	0,378	2,097		2,063
LKGS	3	AC05207c	MB	0,386	2,961	0,384	2,858
KG	6	AB05891c	MB	0,232	2,593	0,190	2,576
KG	6	AB058a1c	MB	0,094	2,286	0,353	2,355

Zwischen den Methoden zeigte sich eine Differenz der Mittelwerte von 0,028s in der Kontrollgruppe (Mediandifferenz 0,045s). In der LKGS-Gruppe betrug die Differenz auf Signallevel im Mittel 0,058s (Mediandifferenz 0,078s).

Die deskriptive Statistik der absoluten Zeitmessungsdifferenz zwischen den Vokalisationen vom **Typ 3B/MB** ist in Tabelle 16 berichtet (Einzelwerte im Anhang A 2). Hier ist nicht zwischen den Gruppen unterschieden worden, da es in der LKGS-Gruppe nur sechs Vokalisationen gab.

Tabelle 16: Deskriptive Statistik absolute Zeitmessungsdifferenzen [s] für die Gesamtlänge der Vokalisationen des Typs 3B/MB zwischen SAM- und MM-Methode für beide Gruppen

Diff_L		Statistik	Std.-Fehler
Mittelwert		,03679	,009462
95% Konfidenzintervall des Mittelwertes	Untergrenze	,01673	
	Obergrenze	,05685	
5% getrimmtes Mittel		,03899	
Median		,03443	
Varianz		,002	
Std.-Abweichung		,039014	
Minimum		-,069	
Maximum		,103	

### 3.3.2.4 Temporale Eigenschaften einfachsegmentierte (1S) Strukturen

Abbildung 52 zeigt die Verteilung der Gesamtlängen im Gruppen- und Methodenvergleich für die einfach segmentierten Melodiestruckturen (Signallevel) in Form eines Boxplotdiagramms. Die zugehörigen deskriptiven Statistiken sind in Tabelle 17 (SAM) und in Tabelle 18 (MM) berichtet. In Tabelle 19 sind alle mit der semi-automatischen Methode (SAM) und in Tabelle 20 alle mit der manuellen Methode (MM) ermittelten Messwerte der Vokalisationen (Säuglingsschreie) mit einfach segmentierter Melodiekontur pro Säugling und Gruppe berichtet. Für die SAM-Messungen sind auch hier wieder die der jeweiligen Vokalisation vorangehenden bzw. nachfolgenden Inspirationslängen (soweit eindeutig vermessenbar) zu Referenzzwecken mit aufgeführt. In den Tabellen sind nicht nur die Gesamtlängen, sondern zusätzlich auch die Zeitdauern der beiden Melodiebögen (B1 und B2) sowie der dazwischenliegenden Oszillationspause (p12) berichtet.

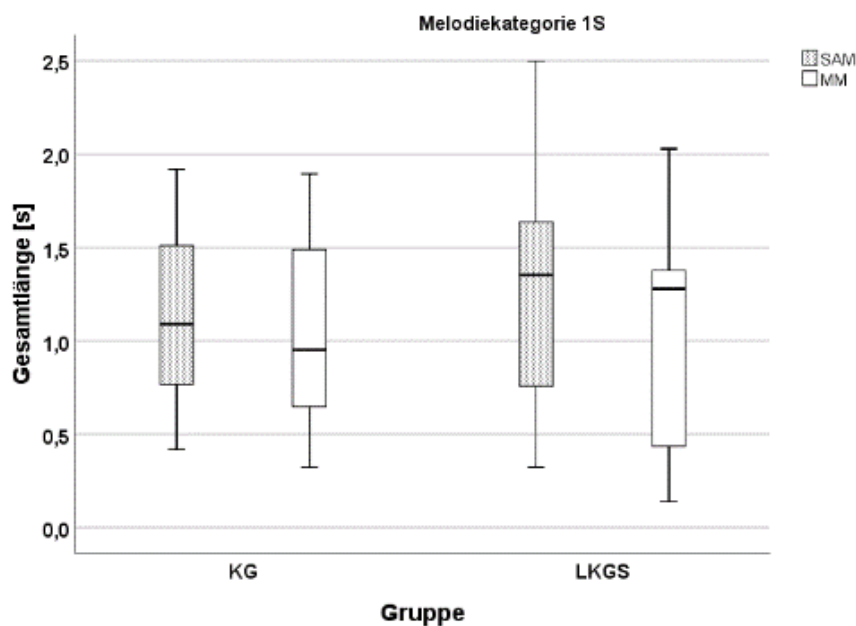


Abbildung 52: Boxplotdiagramm: Vergleich der Gesamtlängen (Lges [s]) von 1S-Strukturen im Gruppen- und Methodenvergleich



Tabelle 17: Deskriptive Statistik für die Gesamtlängen (Lges [s]); SAM-Methode im Gruppenvergleich für 1S- Strukturen

	Gruppe		Statistik	Std.-Fehler	
SAM	KG	Mittelwert	1,13276	,091760	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	,94246	
			Obergrenze	1,32305	
		5% getrimmtes Mittel	1,12979		
		Median	1,08900		
		Varianz	,194		
		Std.-Abweichung	,440067		
		Minimum	,419		
		Maximum	1,919		
	LKGS	Mittelwert	1,30504	,159201	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	,96755	
			Obergrenze	1,64253	
		5% getrimmtes Mittel	1,29328		
		Median	1,35309		
		Varianz	,431		
		Std.-Abweichung	,656401		
		Minimum	,324		
		Maximum	2,498		

Tabelle 18: Deskriptive Statistik für die Gesamtlängen (Lges [s]); MM-Methode im Gruppenvergleich für 1S- Strukturen

	Gruppe		Statistik	Std.-Fehler	
MM	KG	Mittelwert	1,05420	,101043	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	,84465	
			Obergrenze	1,26375	
		5% getrimmtes Mittel	1,04938		
		Median	,95250		
		Varianz	,235		
		Std.-Abweichung	,484586		
		Minimum	,323		
		Maximum	1,894		
	LKGS	Mittelwert	1,02209	,148529	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	,70722	
			Obergrenze	1,33695	
		5% getrimmtes Mittel	1,01528		
		Median	1,27875		
		Varianz	,375		
		Std.-Abweichung	,612401		
		Minimum	,139		
		Maximum	2,028		

Tabelle 19: Messwerte der Vokalisationslängen (Lges [s]) mit der semi-automatischen Methode (SAM) für Strukturen des Typs 1S (I1 und I2 = Inspirationslängen; B1 und B2 = Bogenlängen; p12 = Oszillationspausenlänge; alle in [s])

Gruppe	Pb	Signal	I1	SAM	SAM B1	SAM p12	SAM B2	I2
LKGS	1	AA08213c	0,748	0,324	0,196	0,070	0,058	0,229
LKGS	1	AA08221m	0,100	0,758	0,283	0,147	0,328	0,267
LKGS	1	AA08227c	0,281	1,325	0,188	0,215	0,922	0,347
LKGS	1	AA08217n	0,372	0,442	0,197	0,107	0,138	
LKGS	1	AA08220m	0,142	0,757	0,484	0,062	0,211	
LKGS	2	AB05302n	0,491	1,412	0,390	0,868	0,155	0,221
LKGS	2	AB05306m	0,378	1,253	0,642	0,480	0,130	0,366
LKGS	2	AB05314n	0,806	2,383	0,222	1,887	0,274	1,162
LKGS	3	AC05202c	0,240	1,418	0,964	0,148	0,306	0,294
LKGS	3	AC05216c	0,282	1,353	0,853	0,211	0,289	0,261
LKGS	3	AC05217c	0,280	1,572	1,084	0,207	0,281	0,317
LKGS	3	AC05218c	0,347	0,725	0,215	0,106	0,404	0,051
LKGS	3	AC05219c	0,264	1,823	1,361	0,182	0,280	0,105
LKGS	3	AC05227c	0,335	1,635	0,414	0,063	1,157	0,448
LKGS	3	AC05228c	0,464	2,498	1,838	0,137	0,523	0,413
LKGS	3	AC05235c	0,432	0,485	0,342	0,104	0,039	0,252
LKGS	4	AD06210c	0,359	2,022	1,486	0,132	0,404	0,347
KG	5	AA05671c	0,191	1,389	0,637	0,394	0,358	0,584
KG	6	AB05803m	1,049	0,883	0,280	0,347	0,256	0,377
KG	6	AB05812c	0,221	1,337	0,500	0,323	0,513	0,429
KG	6	AB05834m	0,121	1,609	1,210	0,307	0,092	0,527
KG	6	AB05846c	0,069	0,419	0,327	0,064	0,028	0,335
KG	6	AB05848c	0,148	0,630	0,532	0,066	0,032	0,311
KG	6	AB05860n	0,196	0,465	0,103	0,049	0,313	0,236
KG	6	AB05862n	0,074	0,728	0,283	0,296	0,149	0,268
KG	6	AB05873c	0,233	1,919	0,493	0,155	1,271	0,440
KG	6	AB05875c	0,397	1,564	0,670	0,276	0,618	0,900
KG	6	AB05889m	0,104	0,807	0,703	0,039	0,065	0,194
KG	6	AB05890c	0,072	0,857	0,761	0,032	0,064	0,214
KG	6	AB058b2c	0,277	1,732	0,591	0,312	0,828	0,453
KG	6	AB058b6c	0,492	0,711	0,587	0,047	0,077	0,185
KG	6	AB05826m	0,397	1,267	0,566	0,129	0,572	
KG	6	AB058b5c	0,119	1,455	0,726	0,257	0,472	0,267
KG	6	AB05839c	0,170	1,631	1,060	0,261	0,310	0,488
KG	8	AD06910m	0,329	0,934	0,807	0,083	0,044	0,319
KG	8	AD06968c	0,267	1,699	1,315	0,251	0,133	0,414
KG	8	AD06973m	0,393	0,916	0,417	0,355	0,144	1,008
KG	8	AD06995m	0,241	0,725	0,366	0,209	0,150	0,188
KG	8	AD069a0c	0,211	1,089	0,246	0,213	0,630	0,192
KG	8	AD069c9m	0,174	1,289	0,146	0,123	1,020	0,254

Tabelle 20: Messwerte der Vokalisationslängen (Lges [s]) mit der manuellen Methode (MM) für Strukturen des Typs 1S (I1 und I2 = Inspirationslängen; B1 und B2 = Bogenlängen; p12 = Oszillationspausenlänge; alle in [s])

Gruppe	Pb	Signal	MM	MM B1	MM p12	MM B2
LKGS	1	AA08213c	0,296	0,139	0,109	0,049
LKGS	1	AA08221m	0,615	0,143	0,184	0,289
LKGS	1	AA08227c	1,286	0,068	0,346	0,871
LKGS	1	AA08217n	0,424	0,109	0,274	0,041
LKGS	1	AA08220m	0,409	0,409		
LKGS	2	AB05302n	1,279	0,285	0,979	0,015
LKGS	2	AB05306m	0,435	0,435		
LKGS	2	AB05314n	0,139	0,139		
LKGS	3	AC05202c	1,343	0,893	0,169	0,281
LKGS	3	AC05216c	1,317	0,825	0,255	0,237
LKGS	3	AC05217c	1,377	1,059	0,240	0,078
LKGS	3	AC05218c	0,675	0,141	0,201	0,333
LKGS	3	AC05219c	1,674	1,332	0,213	0,129
LKGS	3	AC05227c	1,569	1,569		
LKGS	3	AC05228c	2,028	1,377	0,138	0,513
LKGS	3	AC05235c	0,537	0,051	0,156	0,330
LKGS	4	AD06210c	1,974	1,449	0,147	0,378
KG	5	AA05671c	1,376	0,555	0,499	0,323
KG	6	AB05803m	0,870	0,229	0,416	0,225
KG	6	AB05812c	1,324	0,345	0,533	0,446
KG	6	AB05834m	1,613	1,219	0,330	0,064
KG	6	AB05846c	0,323	0,323		
KG	6	AB05848c	0,521	0,521		
KG	6	AB05860n	0,458	0,098	0,060	0,300
KG	6	AB05862n	0,630	0,165	0,383	0,083
KG	6	AB05873c	1,894	0,510	0,068	1,316
KG	6	AB05875c	1,526	0,773	0,128	0,626
KG	6	AB05889m	0,630	0,630		
KG	6	AB05890c	0,840	0,705	0,075	0,060
KG	6	AB058b2c	1,703	0,514	0,386	0,803
KG	6	AB058b6c	0,690	0,548	0,075	0,068
KG	6	AB05826m	1,223	0,555	0,086	0,581
KG	6	AB058b5c	1,451	0,731	0,229	0,491
KG	6	AB05839c	1,613	1,099	0,203	0,311
KG	8	AD06910m	0,908	0,746	0,135	0,026
KG	8	AD06968c	1,691	1,301	0,218	0,173
KG	8	AD06973m	0,352	0,352		
KG	8	AD06995m	0,667	0,319	0,218	0,131
KG	8	AD069a0c	0,994	0,060	0,319	0,615
KG	8	AD069c9m	0,953	0,285	0,218	0,450

Zwischen den Methoden zeigte sich eine Mittelwertdifferenz von 0,079s in der Kontrollgruppe (Mediandifferenz 0,137s). In der LKGS-Gruppe betrug die Differenz auf Signallevel im Mittel 0,283s (Mediandifferenz 0,074s).

Die Gruppenunterschiede für die temporalen Messgrößen der Teilzeiten der Strukturelemente sind im Boxplotdiagramm (Abbildung 53) dargestellt. Hier wurden nur die Werte der SAM-Methode dargestellt.

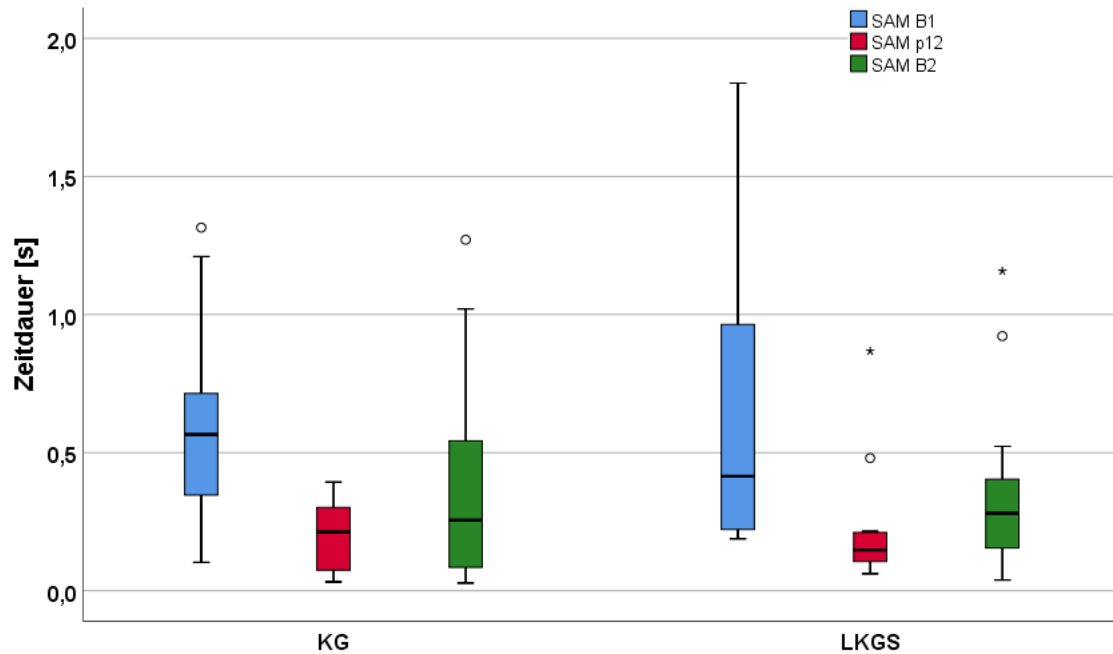


Abbildung 53: Gruppenunterschiede für die temporalen Messgrößen der Teilzeiten einzelner Strukturelemente im Gruppenvergleich (1S, SAM-Methode) in [s]

Die deskriptive Statistik für die Teilzeiten der einzelnen Strukturelemente sind in Tabelle 21 für die SAM-Methode berichtet, für die MM-Methode in Tabelle 22.

Tabelle 21: Deskriptive Statistik für die Teilzeiten der Melodiekonturelemente des Strukturtyps 1S (SAM-Methode) in [s]

	Gruppe		Statistik	Std.-Fehler	
SAM B1	KG	Mittelwert	,57941	,065312	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	,44396	
			Obergrenze	,71485	
		5% getrimmtes Mittel	,56546		
		Median	,56600		
		Varianz	,098		
		Std.-Abweichung	,313224		
		Minimum	,103		
		Maximum	1,315		
	LKGS	Mittelwert	,65646	,126035	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	,38927	
			Obergrenze	,92364	
		5% getrimmtes Mittel	,61684		
		Median	,41449		
		Varianz	,270		
		Std.-Abweichung	,519656		
		Minimum	,188		
		Maximum	1,838		
SAM p12	KG	Mittelwert	,19946	,024775	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	,14808	
			Obergrenze	,25084	
		5% getrimmtes Mittel	,19819		
		Median	,21300		
		Varianz	,014		
		Std.-Abweichung	,118815		
		Minimum	,032		
		Maximum	,394		
	LKGS	Mittelwert	,30149	,109985	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	,06834	
			Obergrenze	,53465	
		5% getrimmtes Mittel	,22673		
		Median	,14728		
		Varianz	,206		
		Std.-Abweichung	,453479		
		Minimum	,062		
		Maximum	1,887		
SAM B2	KG	Mittelwert	,35389	,070995	
		Untergrenze	,20665		

Gruppe		Statistik	Std.-Fehler		
		95% Konfidenzintervall des Mittelwerts	Obergrenze	,50112	
		5% getrimmtes Mittel		,32285	
		Median		,25600	
		Varianz		,116	
		Std.-Abweichung		,340482	
		Minimum		,028	
		Maximum		1,271	
		LKGS	Mittelwert		,34709
	95% Konfidenzintervall des Mittelwerts		Untergrenze	,19670	
			Obergrenze	,49747	
	5% getrimmtes Mittel		,31918		
	Median		,28068		
	Varianz		,086		
	Std.-Abweichung		,292494		
Minimum			,039		
Maximum		1,157			

Tabelle 22: Deskriptive Statistik für die Teilzeiten der Melodiekonturelemente des Strukturtyps 1S (MM-Methode) in [s]

Gruppe		Statistik	Std.-Fehler		
MM B1	KG	Mittelwert		,56605	,082322
		95% Konfidenzintervall des Mittelwerts	Untergrenze	,39310	
			Obergrenze	,73900	
		5% getrimmtes Mittel		,55332	
		Median		,54750	
		Varianz		,129	
		Std.-Abweichung		,358834	
		Minimum		,060	
		Maximum		1,301	
	LKGS	Mittelwert		,60534	,155348
		95% Konfidenzintervall des Mittelwerts	Untergrenze	,26686	
			Obergrenze	,94381	
		5% getrimmtes Mittel		,58926	
		Median		,28500	
		Varianz		,314	
		Std.-Abweichung		,560114	
		Minimum		,051	
		Maximum		1,449	

	Gruppe		Statistik	Std.-Fehler	
MM p12	KG	Mittelwert		,24079	,034478
		95% Konfidenzintervall des Mittelwerts	Untergrenze	,16835	
			Obergrenze	,31322	
		5% getrimmtes Mittel		,23463	
		Median		,21750	
		Varianz		,023	
		Std.-Abweichung		,150284	
		Minimum		,060	
		Maximum		,533	
	LKGS	Mittelwert		,26232	,062316
		95% Konfidenzintervall des Mittelwerts	Untergrenze	,12655	
			Obergrenze	,39810	
		5% getrimmtes Mittel		,23105	
		Median		,20100	
		Varianz		,050	
		Std.-Abweichung		,224684	
		Minimum		,109	
		Maximum		,979	
MM B2	KG	Mittelwert		,37322	,074326
		95% Konfidenzintervall des Mittelwerts	Untergrenze	,21707	
			Obergrenze	,52938	
		5% getrimmtes Mittel		,34011	
		Median		,31125	
		Varianz		,105	
		Std.-Abweichung		,323980	
		Minimum		,026	
		Maximum		1,316	
	LKGS	Mittelwert		,27265	,065267
		95% Konfidenzintervall des Mittelwerts	Untergrenze	,13044	
			Obergrenze	,41485	
		5% getrimmtes Mittel		,25370	
		Median		,28125	
		Varianz		,055	
		Std.-Abweichung		,235323	
		Minimum		,015	
		Maximum		,871	

In Abbildung 54 sind die Verteilungseigenschaften aller temporalen Größen, ermittelt über beide Methoden, vergleichend dargestellt.

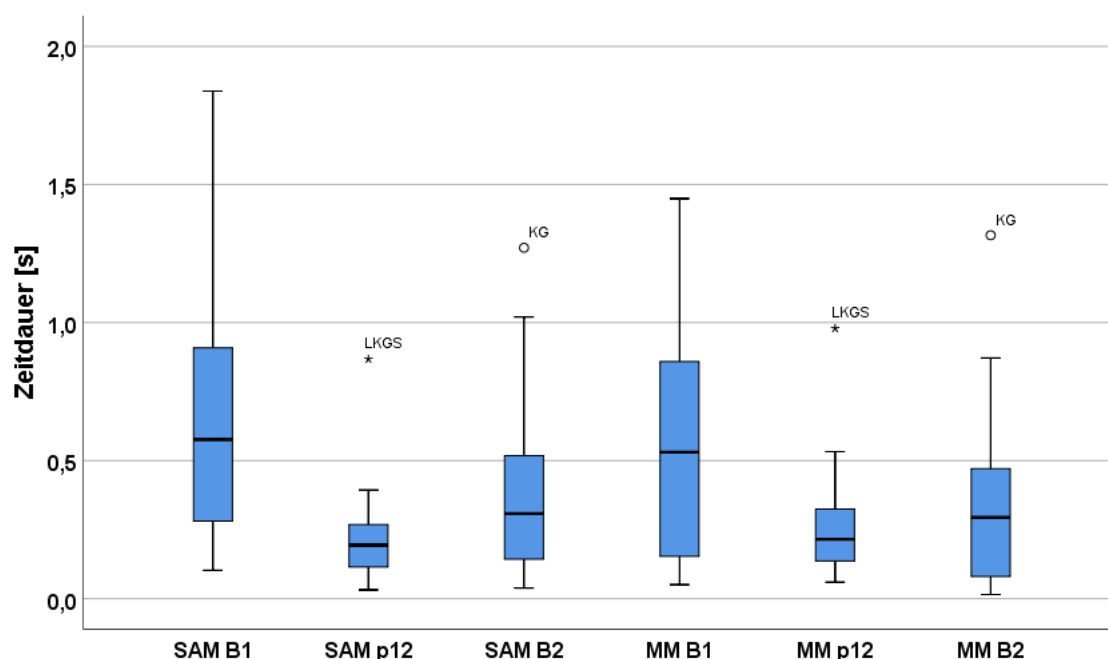


Abbildung 54: Boxplotdiagramm: Gesamtlängen (Lges [s]) aller temporalen Messgrößen für Vokalisationen mit dem melodischen Strukturtyp 1S aus beiden Gruppen

Man erkennt nur geringfügige Unterschiede zwischen den Teillängen der Anfangsbögen (B1), Pausen (p) und nachfolgenden Bögen (B2) zwischen der semi-automatischen, zeitsignalbasierten (SAM) und melodiebasierten (MM) Messmethode. In acht Vokalisationen war mit der MM-Methode keine hinreichend genaue Zeitmessung möglich (vgl. Tabelle 20), so dass die Boxplots nur die Vokalisationen (N=32) mit vollständigen Messungen enthalten.

Die deskriptiven Statistiken der absoluten Zeitmessungsdifferenz zwischen den einzelnen Melodiekonturelementen sind in Tabelle 23 bis Tabelle 26 berichtet (Einzelwerte in Anhang A 2).



Tabelle 23: Deskriptive Statistik: Zeitmessungsdifferenzen [s] für die Gesamtlänge der Vokalisationen zwischen SAM- und MM-Methode für beide Gruppen

	Gruppe		Statistik	Std.-Fehler	
ABS_DIFFLGs	KG	Mittelwert	,0786	,02709	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	,0224	
			Obergrenze	,1347	
		5% getrimmtes Mittel	,0577		
		Median	,0261		
		Varianz	,017		
		Std.-Abweichung	,12993		
		Minimum	,00		
		Maximum	,56		
	LKGS	Mittelwert	,2829	,13295	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	,0011	
			Obergrenze	,5648	
		5% getrimmtes Mittel	,1926		
		Median	,0752		
		Varianz	,300		
		Std.-Abweichung	,54817		
		Minimum	-,05		
		Maximum	2,24		

Tabelle 24: Deskriptive Statistik: Zeitmessungsdifferenzen [s] für den Melodiebogen vor der Pause (B1) der Vokalisationen zwischen SAM- und MM-Methode für beide Gruppen

	Gruppe		Statistik	Std.-Fehler	
ABS_DIFFB1	KG	Mittelwert		,0595	,01080
		95% Konfidenzintervall des Mittelwerts	Untergrenze	,0371	
			Obergrenze	,0819	
		5% getrimmtes Mittel		,0558	
		Median		,0509	
		Varianz		,003	
		Std.-Abweichung		,05180	
		Minimum		,00	
		Maximum		,19	
	LKGS	Mittelwert		,1793	,06669
		95% Konfidenzintervall des Mittelwerts	Untergrenze	,0379	
			Obergrenze	,3207	
		5% getrimmtes Mittel		,1337	
		Median		,0834	
		Varianz		,076	
		Std.-Abweichung		,27496	
		Minimum		,03	
		Maximum		1,15	

Tabelle 25: Deskriptive Statistik: Zeitmessungsdifferenzen [s] für die Pausenlänge der Vokalisationen zwischen SAM- und MM-Methode für beide Gruppen

	Gruppe		Statistik	Std.-Fehler	
ABS_DIFFP	KG	Mittelwert		,0689	,01154
		95% Konfidenzintervall des Mittelwerts	Untergrenze	,0447	
			Obergrenze	,0932	
		5% getrimmtes Mittel		,0645	
		Median		,0585	
		Varianz		,003	
		Std.-Abweichung		,05030	
		Minimum		,01	
		Maximum		,21	
	LKGS	Mittelwert		,0598	,01392
		95% Konfidenzintervall des Mittelwerts	Untergrenze	,0294	
			Obergrenze	,0901	
		5% getrimmtes Mittel		,0571	
		Median		,0389	
		Varianz		,003	
		Std.-Abweichung		,05021	
		Minimum		,00	
		Maximum		,17	

Tabelle 26: Deskriptive Statistik: Zeitmessungsdifferenzen [s] für den Melodienbogen nach der Pause (B2) der Vokalisationen zwischen SAM- und MM-Methode für beide Gruppen

	Gruppe		Statistik	Std.-Fehler	
ABS_DIFFB2	KG	Mittelwert	,0540	,02900	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	-,0070	
			Obergrenze	,1149	
		5% getrimmtes Mittel	,0282		
		Median	,0192		
		Varianz	,016		
		Std.-Abweichung	,12641		
		Minimum	,00		
		Maximum	,57		
		LKGS	Mittelwert	,0896	,02355
	95% Konfidenzintervall des Mittelwerts		Untergrenze	,0383	
			Obergrenze	,1409	
	5% getrimmtes Mittel		,0829		
	Median		,0522		
	Varianz		,007		
	Std.-Abweichung		,08491		
	Minimum		,01		
	Maximum		,29		

### 3.3.2.5 Temporale Eigenschaften zweifachfachsegmentierter Melodien (2S Strukturen)

Abbildung 55 zeigt formal die Verteilung der Gesamtlängen im Gruppen- und Methodenvergleich für die zweifach segmentierten Melodiestrukturen (Signallevel) in Form eines Boxplotdiagramms – allerdings gab es nur einen Säugling in der KG-Gruppe, der Laute dieses Strukturtyps erzeugte. Die zugehörige deskriptive Statistik ist in Tabelle 27 (SAM) und Tabelle 28 (MM) berichtet. Auch hier gilt zu beachten, dass es in der KG-Gruppe nur zwei Vokalisationen gab. In Tabelle 29 sind die mit der semi-automatischen Methode (SAM) und in Tabelle 30 alle mit der manuellen Methode (MM) ermittelten Messwerte der Vokalisationen (Säuglingsschreie) mit zweifach segmentierter Melodiekontur pro Säugling und Gruppe berichtet. Für die SAM-Messungen sind auch hier wieder die der jeweiligen Vokalisation vorangehenden bzw. nachfolgenden Inspirationslängen (soweit eindeutig vermessbar) mit aufgeführt.

In den Tabellen sind nicht nur die Gesamtlängen, sondern zusätzlich auch die Zeitauern der Melodiebögen (B1, B2 und B3) sowie der dazwischenliegenden Oszillationspausen (p12 und p23) berichtet.

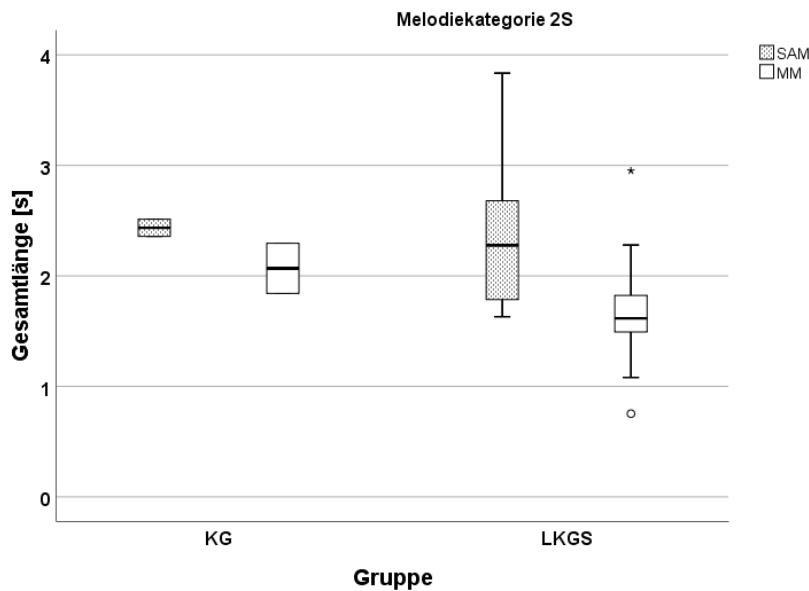


Abbildung 55: Boxplotdiagramm: Vergleich der Gesamtlängen (Lges [s]) von 2S-Strukturen im Gruppen- und Methodenvergleich

Tabelle 27: Deskriptive Statistik für die Gesamtlängen (Lges [s]); SAM-Methode im Gruppenvergleich für 2S-Strukturen

Gruppe		Statistik	Std.-Fehler		
SAM	KG	Mittelwert	2,43500	,078000	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	1,44392	
			Obergrenze	3,42608	
		5% getrimmtes Mittel		.	
		Median	2,43500		
		Varianz	,012		
		Std.-Abweichung	,110309		
	LKGS	Mittelwert	2,30746	,178640	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	1,91824	
			Obergrenze	2,69669	
		5% getrimmtes Mittel	2,26028		
		Median	2,27700		
		Varianz	,415		
		Std.-Abweichung	,644097		
Minimum	1,630				
Maximum	3,835				

Tabelle 28: Deskriptive Statistik für die Gesamtlängen (Lges [s]); MM-Methode im Gruppenvergleich für 2S-Strukturen

	Gruppe		Statistik	Std.-Fehler	
MM	KG	Mittelwert	2,06812	,226875	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	-,81460	
			Obergrenze	4,95085	
		5% getrimmtes Mittel		.	
		Median	2,06812		
		Varianz	,103		
		Std.-Abweichung	,320850		
	LKGS	Mittelwert	1,68896	,147909	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	1,36670	
			Obergrenze	2,01123	
		5% getrimmtes Mittel	1,67086		
		Median	1,61625		
		Varianz	,284		
		Std.-Abweichung	,533294		
Minimum	,753				
Maximum	2,951				

Tabelle 29: Messwerte aller Vokalisationslängen (Lges [s]) mit der semi-automatischen Methode (SAM) für Strukturen des Typs 2S (I1 und I2 = Inspirationslängen; B1-B3 = Bogenlängen; p12, p23= Oszillationspausenlängen; alle in [s])

Gruppe	Pb	Signal	I1	SAM	SAM B1	SAM p12	SAM B2	SAM p23	SAM B3	I2
LKGS	1	AA08207m	0,904	1,767	0,282	0,568	0,362	0,268	0,287	0,653
LKGS	1	AA08208c	0,691	1,855	0,465	0,085	0,515	0,044	0,747	0,485
LKGS	1	AA08209n	0,551	2,306	0,338	0,180	0,777	0,227	0,784	0,733
LKGS	1	AA08212c	1,142	2,976	0,670	0,358	0,572	1,354	0,023	0,496
LKGS	1	AA08214c	0,264	2,356	0,082	0,160	0,948	1,071	0,096	0,270
LKGS	1	AA08222m	0,279	2,679	1,396	0,208	0,625	0,327	0,123	0,571
LKGS	2	AB05308m	0,574	3,835	1,001	0,369	1,943	0,380	0,142	0,328
LKGS	2	AB05309m	0,502	2,277	0,604	0,550	0,533	0,441	0,149	0,391
LKGS	2	AB05318c	0,365	2,868	0,764	0,863	0,905	0,201	0,135	0,289
LKGS	3	AC05223c	0,297	1,656	0,747	0,081	0,278	0,091	0,459	0,356
LKGS	3	AC05226c	0,352	2,006	0,808	0,132	0,550	0,100	0,416	0,313
LKGS	3	AC05230c	0,319	1,786	0,692	0,187	0,299	0,481	0,127	0,273
LKGS	3	AC05231c	0,304	1,630	0,813	0,159	0,377	0,098	0,183	0,272
KG	6	AB05819c	0,740	2,513	0,933	0,190	0,816	0,486	0,088	0,486
KG	6	AB05864c	0,136	2,357	1,479	0,328	0,363	0,075	0,112	0,552

Tabelle 30: Messwerte aller Vokalisationslängen (Lges [s]) mit der manuellen Methode (MM) für Strukturen des Typs 2S (I1 und I2 = Inspirationslängen; B1-B3 = Bogenlängen; p12, p23 = Oszillationspausenlängen; alle in [s])

Gruppe	Pb	Signal	MM	MM B1	MM p12	MM B2	MM p23	MM B3
LKGS	1	AA08207m	1,616	0,135	0,611	0,311	0,345	0,214
LKGS	1	AA08208c	1,823	0,338	0,236	0,458	0,086	0,705
LKGS	1	AA08209n	1,624	0,623	0,634	0,368		
LKGS	1	AA08212c	1,493	0,611	0,413	0,469		
LKGS	1	AA08214c	0,753	0,753				
LKGS	1	AA08222m	1,080	1,080				
LKGS	2	AB05308m	2,951	0,814	0,559	1,579		
LKGS	2	AB05309m	1,444	0,536	0,675	0,233		
LKGS	2	AB05318c	2,280	0,615	0,870	0,795		
LKGS	3	AC05223c	1,590	0,729	0,096	0,222	0,207	0,336
LKGS	3	AC05226c	1,962	0,804	0,105	0,372	0,615	0,066
LKGS	3	AC05230c	1,746	0,666	0,213	0,255	0,531	0,081
LKGS	3	AC05231c	1,596	0,789	0,177	0,333	0,171	0,126
KG	6	AB05819c	1,841	0,885	0,199	0,758		
KG	6	AB05864c	2,295	1,425	0,334	0,334	0,113	0,090

Zwischen den Methoden zeigte sich eine Differenz der Mittelwerte von 0,367s in der Kontrollgruppe (Mediandifferenz 0,367s). In der LKGS-Gruppe betrug die Differenz auf Signallevel im Mittel 0,619s (Mediandifferenz 0,661s).

In Tabelle 31 ist die deskriptive Statistik der absoluten Zeitmessungsdifferenzen zwischen den mit der SAM- bzw. MM-Methode ermittelten Werten dargestellt. Tabelle 32 zeigt die die deskriptive Statistik der Zeitmessungsdifferenzen zwischen den Methoden der einzelnen Strukturelemente für die LKGS-Gruppe.

Tabelle 31: Deskriptive Statistik der ermittelten Differenzen [s] zwischen SAM und MM-Methode für Strukturen der Strukturkategorie 2S

	Gruppe		Statistik	Std.-Fehler	
Diff_LG	KG	Mittelwert	,36688	,304875	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	-3,50693	
			Obergrenze	4,24068	
		5% getrimmtes Mittel	.		
		Median	,36688		
		Varianz	,186		
		Std.-Abweichung	,431158		
		Minimum	,062		
		Maximum	,672		

	Gruppe		Statistik	Std.-Fehler	
	LKGS	Mittelwert	,61850	,173134	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	,24128	
			Obergrenze	,99573	
		5% getrimmtes Mittel	,59631		
		Median	,58800		
		Varianz	,390		
		Std.-Abweichung	,624242		
		Minimum	,033		
		Maximum	1,604		

Tabelle 32: Deskriptive Statistik: Zeitmessungsdifferenzen [s] der einzelnen Strukturelemente der LKGS-Gruppe (2S)

	Gruppe		Statistik	Std.-Fehler	
Diff B1	LKGS	Mittelwert	,05764	,025393	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	-,00763	
			Obergrenze	,12292	
		5% getrimmtes Mittel	,05568		
		Median	,02499		
		Varianz	,004		
		Std.-Abweichung	,062199		
		Minimum	,004		
		Maximum	,147		
Diff P12	LKGS	Mittelwert	-,03761	,024606	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	-,10087	
			Obergrenze	,02564	
		5% getrimmtes Mittel	-,03489		
		Median	-,02175		
		Varianz	,004		
		Std.-Abweichung	,060273		
		Minimum	-,151		
		Maximum	,027		
Diff B2	LKGS	Mittelwert	,07161	,021387	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	,01663	
			Obergrenze	,12659	
		5% getrimmtes Mittel	,06724		
		Median	,05318		
		Varianz	,003		
		Std.-Abweichung	,052387		
		Minimum	,044		



	Gruppe		Statistik	Std.-Fehler	
		Maximum	,178		
Diff P23	LKGS	Mittelwert	-,14556	,074637	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	-,33742	
			Obergrenze	,04630	
		5% getrimmtes Mittel	-,13078		
		Median	-,07515		
		Varianz	,033		
		Std.-Abweichung	,182822		
		Minimum	-,515		
		Maximum	-,042		
Diff B3	LKGS	Mittelwert	,11514	,048495	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	-,00952	
			Obergrenze	,23980	
		5% getrimmtes Mittel	,10617		
		Median	,06506		
		Varianz	,014		
		Std.-Abweichung	,118787		
		Minimum	,042		
		Maximum	,350		

### 3.3.2.6 Temporale Eigenschaften mehrfachsegmentierter (3S) Strukturen

Auch dieser Strukturtyp kam selten vor (siehe Tabellen 35, 36 und 37). Abbildung 56 zeigt die Verteilung der Gesamtlängen im Gruppen- und Methodenvergleich für die dreifach segmentierten Melodiestructuren (Signallevel) in Form eines Boxplotdiagramms. Die zugehörige deskriptive Statistik ist in Tabelle 33 (SAM) und Tabelle 34 (MM) berichtet. In den Tabellen 35 und 36 sind alle mit der semi-automatischen Methode (SAM) und in Tabelle 37 alle mit der manuellen Methode (MM) ermittelten Messwerte der Vokalisationen (Säuglingsschreie) mit dreifach segmentierter Melodiekontur pro Säugling und Gruppe berichtet. Für die SAM-Messungen sind auch hier wieder die der jeweiligen Vokalisation vorangehenden bzw. nachfolgenden Inspirationslängen (soweit eindeutig vermessbar) mit aufgeführt. In den Tabellen sind nicht nur die Gesamtlängen, sondern zusätzlich auch die Zeitdauern der Melodiebögen (B1, B2, B3 und B4) sowie der dazwischenliegenden Oszillationspausen (p12, p23 und p34) berichtet.

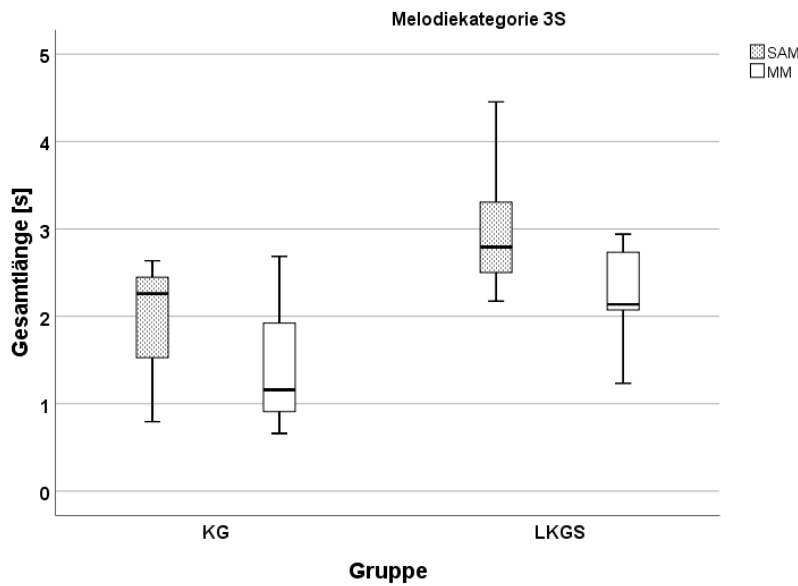


Abbildung 56: Boxplotdiagramm: Vergleich der Gesamtlängen (Lges [s]) von 3S-Strukturen im Gruppen- und Methodenvergleich

Tabelle 33: Deskriptive Statistik für die Gesamtlängen (Lges [s]); SAM-Methode im Gruppenvergleich für 3S-Strukturen

Gruppe		Statistik	Std.-Fehler		
SAM	KG	Mittelwert	1,89657	,561580	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	-,51972	
			Obergrenze	4,31285	
		5% getrimmtes Mittel		.	
		Median	2,25900		
		Varianz	,946		
		Std.-Abweichung	,972684		
		Minimum	,795		
		Maximum	2,636		
	LKGS	Mittelwert	3,04489	,398174	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	1,93938	
			Obergrenze	4,15039	
		5% getrimmtes Mittel	3,01509		
		Median	2,79264		
		Varianz	,793		
Std.-Abweichung		,890343			
Minimum		2,173			
Maximum	4,453				

Tabelle 34: Deskriptive Statistik für die Gesamtlängen (Lges [s]); MM-Methode im Gruppenvergleich für 2S-Strukturen

	Gruppe		Statistik	Std.-Fehler	
MM	KG	Mittelwert	1,50125	,609135	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	-1,11965	
			Obergrenze	4,12215	
		5% getrimmtes Mittel		.	
		Median	1,15875		
		Varianz	1,113		
		Std.-Abweichung	1,055053		
		Minimum	,660		
		Maximum	2,685		
	LKGS	Mittelwert	2,22330	,298551	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	1,39439	
			Obergrenze	3,05221	
		5% getrimmtes Mittel	2,23846		
		Median	2,13600		
		Varianz	,446		
		Std.-Abweichung	,667579		
		Minimum	1,234		
		Maximum	2,940		

Tabelle 35: Messwerte der Vokalisationslängen (Lges [s]) mit der semi-automatischen Methode (SAM) ermittelt für Strukturen des Typs 3S (I1 und I2 = Inspirationslängen; alle in [s])

Gruppe	Pb	Signal	I1	SAM	I2
LKGS	1	AA08218c	0,293	2,500	0,396
LKGS	2	AB05307n	0,289	3,305	0,400
LKGS	2	AB05319c	0,274	4,453	0,415
LKGS	3	AC05220c	0,399	2,173	0,344
LKGS	3	AC05229c	0,376	2,793	0,276
KG	6	AB05814u	0,129	2,636	0,395
KG	6	AB05877m	0,246	0,795	0,317
KG	8	AD069d2c	0,294	2,259	0,251

Tabelle 36: Messwerte der Bogenlängen und Oszillationspausenlängen mit der semi-automatischen Methode (SAM) ermittelt für Strukturen des Typs 3S (B1-B4 = Bogenlängen; p12, p23, p34 = Oszillationspausenlängen; alle in [s])

Gruppe	Pb	Signal	SAM B1	SAM p12	SAM B2	SAM p23	SAM B3	SAM p34	SAM B4
LKGS	1	AA08218c	0,833	0,290	0,296	0,357	0,412	0,192	0,120
LKGS	2	AB05307n	0,307	0,642	0,479	0,892	0,358	0,500	0,128
LKGS	2	AB05319c	0,132	0,107	0,897	0,432	1,818	0,973	0,094
LKGS	3	AC05220c	0,927	0,188	0,275	0,078	0,364	0,205	0,136
LKGS	3	AC05229c	1,400	0,182	0,095	0,122	0,672	0,149	0,172
KG	6	AB05814u	0,454	0,493	0,129	0,556	0,174	0,322	0,508
KG	6	AB05877m	0,008	0,033	0,432	0,032	0,042	0,074	0,174
KG	8	AD069d2c	0,264	0,133	1,322	0,119	0,218	0,091	0,112

Tabelle 37: Messwerte der Vokalisationslängen (Lges [s]) mit der manuellen Methode (MM) ermittelt für Strukturen des Typs 3S (I1 und I2 = Inspirationslängen; B1-B4 = Bogenlängen; p12, p23, p34 = Oszillationspausenlängen; alle in [s])

Gruppe	Pb	Signal	MM	MM B1	MM p12	MM B2	MM p23	MM B3	MM p34	MM B4
LKGS	1	AA08218c	2,074	0,776	0,379	0,158	0,469	0,293		
LKGS	2	AB05307n	1,234	0,161	0,795	0,278				
LKGS	2	AB05319c	2,940	0,803	0,630	0,514	0,570	0,424		
LKGS	3	AC05220c	2,136	0,900	0,210	0,255	0,123	0,315	0,258	0,075
LKGS	3	AC05229c	2,733	2,733						
KG	6	AB05814u	2,685	0,375	0,563	0,105	0,454	0,304	0,311	0,574
KG	6	AB05877m	0,660	0,660						
KG	8	AD069d2c	1,159	1,159						

Zwischen den Methoden zeigte sich eine Differenz der Mittelwerte von 0,395s in der Kontrollgruppe (Mediandifferenz 1,1s). In der LKGS-Gruppe betrug die Differenz auf Signallevel im Mittel 0,822s (Mediandifferenz 0,657s).

### 3.3.2.7 Temporale Eigenschaften von KS Strukturen

Abbildung 57 zeigt die Verteilung der Gesamtlängen im Gruppen- und Methodenvergleich für die komplex segmentierten Melodiestrutturen (Signallevel) in Form eines Boxplotdiagramms. Dieser Strukturtyp beschreibt verschiedenen Konturen, die alle durch eine einzige innerlautliche Pause charakterisiert sind. Im Unterschied zum Typ 1S (Melodiebogen-Pause-Melodiebogen) ist vor oder nach der Pause bei diesem Typ KS eine komplexere Struktur vorhanden, also mehr als ein einzelner Melodiebogen. Die zugehörige

deskriptive Statistik ist in Tabelle 38 (SAM) und Tabelle 39 (MM) berichtet. In Tabelle 40 sind alle mit der semi-automatischen Methode (SAM) und Tabelle 41 alle mit der manuellen Methode (MM) ermittelten Messwerte der Vokalisationen (Säuglingsschreie) mit komplex segmentierter Melodiekontur pro Säugling und Gruppe berichtet. Für die SAM-Messungen sind auch hier wieder die der jeweiligen Vokalisation vorangehenden bzw. nachfolgenden Inspirationslängen (soweit eindeutig vermessbar) mit aufgeführt. In den Tabellen sind nicht nur die Gesamtlängen, sondern zusätzlich auch die Zeitdauern der beiden Melodiekonturen (B1 und B2) sowie der dazwischenliegenden Oszillationspause (p12) berichtet. Die deskriptive Statistik für die Gesamtlängen der Oszillationspausen im Methodenvergleich sind in Tabelle 42 berichtet. Bei diesem Strukturtyp der Melodie gibt es eine innermelodische Pause, davor und danach können beliebig viele Bögen folgen. Hier gibt der Wert B1 die Länge der Melodiekontur vor der Pause, B2 die der nach der Pause an.

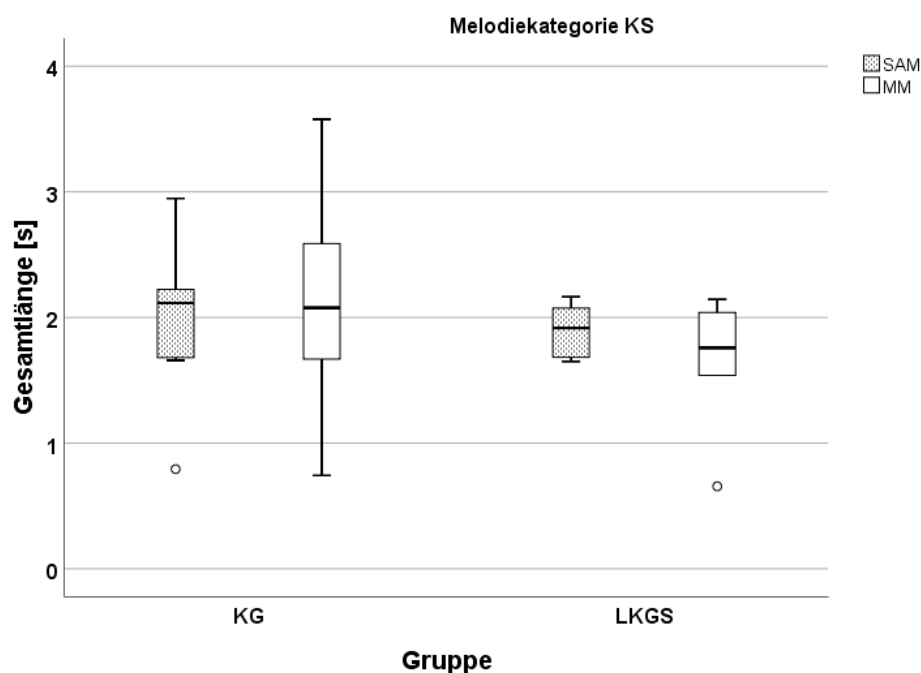


Abbildung 57: Boxplotdiagramm: Vergleich der Gesamtlängen (Lges [s]) von KS-Strukturen im Gruppen- und Methodenvergleich

Tabelle 38: Deskriptive Statistik für die Gesamtlängen (Lges [s]); SAM-Methode im Gruppenvergleich für KS-Strukturen

	Gruppe		Statistik	Std.-Fehler	
SAM	KG	Mittelwert	2,00794	,207294	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	1,52992	
			Obergrenze	2,48596	
		5% getrimmtes Mittel	2,02327		
		Median	2,11487		
		Varianz	,387		
		Std.-Abweichung	,621882		
		Minimum	,793		
		Maximum	2,947		
	LKGS	Mittelwert	1,90099	,084034	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	1,68497	
			Obergrenze	2,11700	
		5% getrimmtes Mittel	1,90031		
		Median	1,91667		
		Varianz	,042		
		Std.-Abweichung	,205840		
		Minimum	1,649		
		Maximum	2,165		

Tabelle 39: Deskriptive Statistik für die Gesamtlängen (Lges [s]); MM-Methode im Gruppenvergleich für KS-Strukturen

	Gruppe		Statistik	Std.-Fehler	
MM	KG	Mittelwert	2,08086	,276855	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	1,44243	
			Obergrenze	2,71929	
		5% getrimmtes Mittel	2,07207		
		Median	2,07750		
		Varianz	,690		
		Std.-Abweichung	,830565		
		Minimum	,742		
		Maximum	3,578		
	LKGS	Mittelwert	1,64925	,218992	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	1,08631	
			Obergrenze	2,21219	
		5% getrimmtes Mittel	1,67688		
		Median	1,75763		
		Varianz	,288		
		Std.-Abweichung	,536419		
		Minimum	,656		
		Maximum	2,145		

Tabelle 40: Messwerte der Vokalisationslängen (Lges [s]) mit der semi-automatischen Methode (SAM) ermittelt für Strukturen des Typs KS (I1 und I2 = Inspirationslängen; B1 und B2 = Bogenlängen; p12 = Oszillationspausenlänge; alle in [s])

Gruppe	Pb	Signal	I1	SAM	SAM B1	SAM p12	SAM B2	I2
LKGS	1	AA08225c	0.277	1.683	1.023	0.116	0.544	0.607
LKGS	2	AB05303m	0.252	1.902	1.495	0.235	0.172	0.250
LKGS	3	AC05212c	0.234	2.075	1.800	0.132	0.143	0.204
LKGS	3	AC05225c	0.298	1.649	1.302	0.120	0.227	0.327
LKGS	3	AC05233c	0.293	1.931	0.973	0.238	0.720	0.288
LKGS	4	AD06211c	0.381	2.165	0.750	0.161	1.254	0.226
KG	5	AA05626c	0.619	0.793	0.208	0.225	0.360	0.450
KG	5	AA05663c	0.189	1.835	0.186	0.158	1.491	0.454
KG	5	AA05673c	0.086	2.190	1.220	0.157	0.813	0.476
KG	6	AB05887m	0.137	2.224	1.337	0.050	0.837	0.304
KG	6	AB05827m	0.121	1.680	1.032	0.416	0.232	0.313
KG	6	AB05831c	0.083	2.947	2.349	0.421	0.177	0.262
KG	7	AC05605m	0.077	2.629	1.980	0.450	0.199	0.305
KG	7	AC05636m	0.048	2.115	1.368	0.253	0.493	0.132
KG	8	AD069b2m	0.327	1.659	0.462	0.207	0.990	0.186

Tabelle 41: Messwerte der Vokalisationslängen (Lges [s]) mit der manuellen Methode (MM) ermittelt für Strukturen des Typs KS (I1 und I2 = Inspirationslängen; B1 und B2 = Bogenlängen; p12 = Oszillationspausenlänge; alle in [s])

Gruppe	Pb	Signal	MM	MM B1	MM p12	MM B2
LKGS	1	AA08225c	1.661	1.009	0.135	0.518
LKGS	2	AB05303m	0.656	0.656		
LKGS	3	AC05212c	2.040	1.755	0.215	0.070
LKGS	3	AC05225c	1.539	1.281	0.177	0.081
LKGS	3	AC05233c	1.854	0.957	0.255	0.642
LKGS	4	AD06211c	2.145	0.738	0.171	1.236
KG	5	AA05626c	0.743	0.176	0.266	0.300
KG	5	AA05663c	1.804	0.225	0.075	1.504
KG	5	AA05673c	2.734	1.234	0.086	1.414
KG	6	AB05887m	2.194	1.234	0.146	0.814
KG	6	AB05827m	1.669	0.983	0.488	0.199
KG	6	AB05831c	3.578	2.261	0.480	0.836
KG	7	AC05605m	2.588	1.958	0.514	0.116
KG	7	AC05636m	2.078	1.320	0.315	0.443
KG	8	AD069b2m	1.343	0.165	0.270	0.908

Zwischen den Methoden zeigte sich eine Differenz der Mittelwerte von -0,073s in der Kontrollgruppe (Mediandifferenz 0,037s). In der LKGS-Gruppe betrug die Differenz auf Signalleve im Mittel 0,252s (Mediandifferenz 0,159s).

Die deskriptive Statistik der absoluten Zeitmessungsdifferenz auf Signallevel zwischen der mit SAM- bzw. MM-Methode ermittelten Werte der Vokalisationen vom **Typ KS** ist in der Tabelle 43 berichtet (Einzelwerte Anhang A 2).

Tabelle 42: Deskriptive Statistik: Gesamtlänge der Oszillationspausen [s] im Methodenvergleich der Strukturkategorie KS

		Statistik	Std.-Fehler	
SAM p12	Mittelwert	,22172	,033310	
	95% Konfidenzintervall des Mittelwerts	Untergrenze	,14976	
		Obergrenze	,29368	
	5% getrimmtes Mittel	,21856		
	Median	,18419		
	Varianz	,016		
	Std.-Abweichung	,124634		
	Minimum	,050		
	Maximum	,450		
	MM p12	Mittelwert	,25663	,039044
95% Konfidenzintervall des Mittelwerts		Untergrenze	,17228	
		Obergrenze	,34097	
5% getrimmtes Mittel		,25243		
Median		,23500		
Varianz		,021		
Std.-Abweichung		,146089		
Minimum		,075		
Maximum		,514		



Tabelle 43: Deskriptive Statistik der ermittelten Differenzen [s] zwischen SAM und MM-Methode für Strukturen der Strukturkategorie KS

	Gruppe		Statistik	Std.-Fehler	
ABS_DiffLG	KG	Mittelwert	,1880	,08189	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	-,0008	
			Obergrenze	,3769	
		5% getrimmtes Mittel	,1733		
		Median	,0415		
		Varianz	,060		
		Std.-Abweichung	,24568		
		Minimum	,01		
		Maximum	,63		
	LKGS	Mittelwert	,2517	,19938	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	-,2608	
			Obergrenze	,7642	
		5% getrimmtes Mittel	,2094		
		Median	,0560		
		Varianz	,239		
		Std.-Abweichung	,48837		
		Minimum	,02		
		Maximum	1,25		

### 3.3.2.8 Temporale Eigenschaften der unklaren Strukturen „fuzzy class“

Melodiekonturen, deren Struktur stark irregulär war, so dass sie nicht eindeutig einer der definierten Strukturkategorien (vgl Kapitel 2.1) zugeordnet werden konnten, wurden in der „fuzzy class“ zusammengefasst. Die Gesamtlänge der Vokalisationen wurde ebenfalls vermessen.

Abbildung 58 zeigt die Verteilung der Gesamtlängen im Gruppen- und Methodenvergleich für die „fuzzy class“ Melodiestructuren (Signallevel) in Form eines Boxplotdiagramms. Die zugehörige deskriptive Statistik ist in Tabelle 44 (SAM) und Tabelle 45 (MM) berichtet. In Tabelle 46 sind alle Messwerte der Vokalisationen (Säuglingsschreie) „fuzzy class“ pro Säugling und Gruppe berichtet. Für die SAM-Messungen sind auch hier wieder die der jeweiligen Vokalisation vorangehenden bzw. nachfolgenden Inspirationslängen (soweit eindeutig vermessbar) mit aufgeführt.

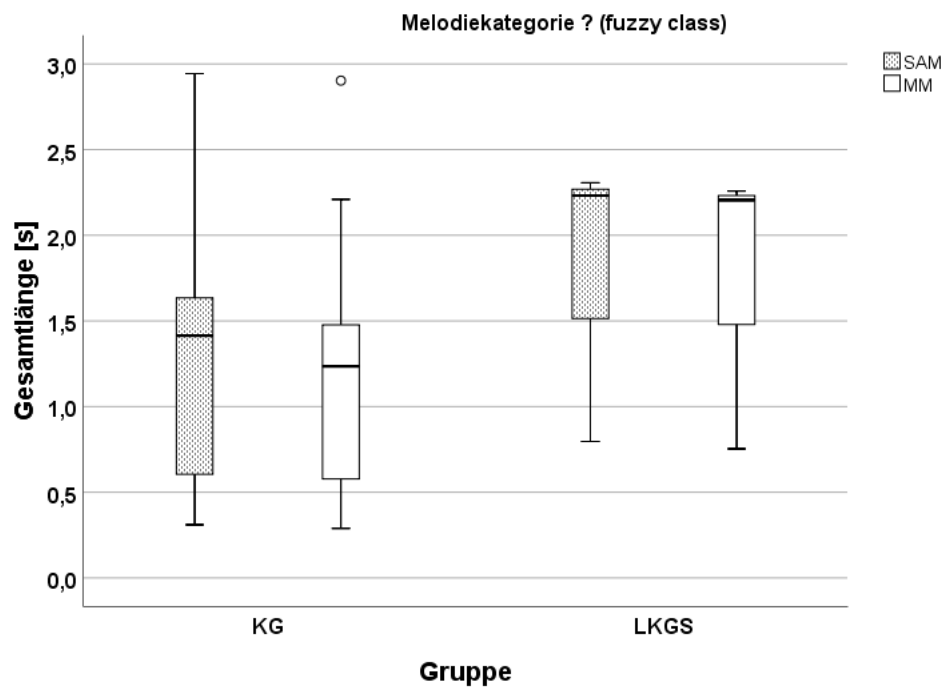


Abbildung 58: Boxplotdiagramm: Vergleich der Gesamtlängen (Lges [s]) von fuzzy class- Strukturen im Gruppen- und Methodenvergleich

Tabelle 44: Deskriptive Statistik für die Gesamtlängen (Lges [s]); SAM-Methode im Gruppenvergleich für fuzzy class-Strukturen

Gruppe		Statistik	Std.-Fehler		
SAM	KG	Mittelwert	1,30360	,196630	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	,88187	
			Obergrenze	1,72532	
		5% getrimmtes Mittel	1,26763		
		Median	1,41403		
		Varianz	,580		
		Std.-Abweichung	,761544		
		Minimum	,311		
		Maximum	2,944		
	LKGS	Mittelwert	1,77806	,491332	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	-,33597	
			Obergrenze	3,89209	
		5% getrimmtes Mittel	.		
		Median	2,23145		
		Varianz	,724		
Std.-Abweichung		,851012			
Minimum		,796			
Maximum	2,306				

Tabelle 45: Deskriptive Statistik für die Gesamtlängen (Lges [s]); MM-Methode im Gruppenvergleich für fuzzy class-Strukturen

	Gruppe		Statistik	Std.-Fehler	
MM	KG	Mittelwert	1,20377	,194043	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	,78759	
			Obergrenze	1,61995	
		5% getrimmtes Mittel	1,16023		
		Median	1,23500		
		Varianz	,565		
		Std.-Abweichung	,751524		
		Minimum	,289		
		Maximum	2,903		
	LKGS	Mittelwert	1,73875	,492733	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	-,38131	
			Obergrenze	3,85881	
		5% getrimmtes Mittel	.		
		Median	2,20500		
		Varianz	,728		
		Std.-Abweichung	,853439		
		Minimum	,754		
		Maximum	2,258		

Tabelle 46: Messwerte der Vokalisationslängen (Lges [s]) mit der semi-automatischen Methode (SAM) und der manuellen Methode (MM) für Strukturen des Typs fuzzy class (I1 und I2 = Inspirationslängen in [s])

Gruppe	Kind	Signal	I1	SAM	I2	MM	Diff_L
LKGS	1	AA08210c	1,009	0,796	0,558	0,754	0,043
LKGS	4	AD06203c	0,342	2,231	0,285	2,205	0,026
LKGS	4	AD06205c	0,570	2,306	0,340	2,258	0,049
KG	5	AA05684m	0,210	0,311	0,232	0,289	0,022
KG	6	AB05849m	0,185	1,548	1,222	1,493	0,055
KG	6	AB058a2c	0,052	1,497	0,218	0,833	0,665
KG	7	AC05606c	0,043	1,990	0,367	1,965	0,025
KG	7	AC05607c	0,057	2,944	0,563	2,903	0,041
KG	7	AC05609c	0,027	0,366	0,244	0,360	0,006
KG	8	AD06912c	0,248	2,227	0,259	2,209	0,018
KG	8	AD06915c	0,632	1,504	0,232	1,443	0,062
KG	8	AD06921c	0,248	1,316	0,355	1,264	0,053
KG	8	AD06939m	0,124	0,345	0,190	0,345	0,000
KG	8	AD06947m	0,534	0,591	0,256	0,570	0,021
KG	8	AD06976c	0,445	1,414	0,818	1,235	0,179
KG	8	AD06996m	0,217	0,618	0,227	0,585	0,033
KG	8	AD069c5m	0,277	1,722	0,748	1,463	0,259
KG	8	AD069d6c	0,225	1,161	0,274	1,103	0,059

Zwischen den Methoden zeigte sich eine Differenz der Mittelwerte von 0,1s in der Kontrollgruppe (Mediandifferenz 0,179s). In der LKGS-Gruppe betrug die Differenz auf Signallevel im Mittel 0,039s (Mediandifferenz 0,026s).

### 3.4 Inspirationslängen

Die Inspirationslängen waren nicht Gegenstand der eigentlichen Analyse, wurden aber durch Anwendung der SAM-Methode quasi automatisch mit ermittelt und wurden hier als Zusatzinformationen für das Projekt und als erste Referenzwerte für geplante nachfolgende Studien mitberichtet. Ein Boxplot der prä- und postphonatorischen Inspirationslängen ist in Abbildung 59, die zugehörige deskriptive Statistik in Tabelle 47 berichtet.

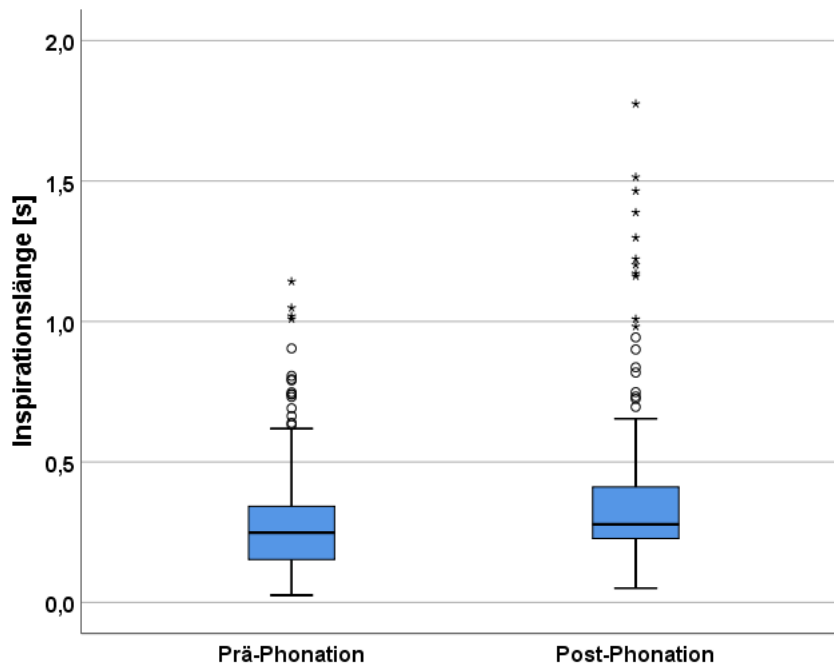


Abbildung 59: Boxplotdiagramm der Inspirationslängen in [s]; Prä- und Post-Phonation (SAM-Methode)

Tabelle 47: Deskriptive Statistik Inspirationslängen [s] (SAM-Methode)

		Statistik	Std.-Fehler	
M_i1	Mittelwert	,277261	,0124114	
	95% Konfidenzintervall des Mittelwerts	Untergrenze	,252814	
		Obergrenze	,301708	
	5% getrimmtes Mittel	,257904		
	Median	,248000		
	Varianz	,038		
	Std.-Abweichung	,1946659		
	Minimum	,0254		
	Maximum	1,1424		
	i2	Mittelwert	,367568	,0160914
95% Konfidenzintervall des Mittelwerts		Untergrenze	,335873	
		Obergrenze	,399263	
5% getrimmtes Mittel		,331882		
Median		,278049		
Varianz		,064		
Std.-Abweichung		,2523832		
Minimum		,0506		
Maximum		1,7747		

## 4 Diskussion

Die Zielsetzung vorliegender Arbeit war projektbedingt komplex und gliederte sich im Wesentlichen in drei Bereiche, wie in Kapitel 1.2 erläutert. Entsprechend wurden die Ergebnisberichte auch in diese drei Bereiche gegliedert. Auch die Diskussion folgt diesem Schema und ist in die folgenden Aufgabenbereiche zusammengefasst: PRAAT-Skriptentwicklung, Testung des Skriptes anhand vorgegebener anonymisierter Patientenaudiofiles und Anwendung der Methode auf konkrete, pseudonymisierte Projektdaten.

### 4.1 Entwicklung und Testung der SAM-Methode

Eine Aufgabenstellung der vorliegenden Arbeit bestand darin, ein, auch für nicht mit der PRAAT-Skriptsprache näher vertraute Benutzer, leicht zu handhabendes Skript für die semi- und teilweise vollautomatische Analyse von aufgezeichneten Lautsequenzen (Sounddateien) zu erarbeiten, zu testen und am Signalanalyselabor des ZVES zu implementieren. Es soll in der Forschung und der Vorsprachlichen Diagnostik dazu eingesetzt werden, aus teilautomatisch annotierten Langserien in PRAAT beliebige gewünschte zeitliche Signalbereiche zu extrahieren und in diesen bestimmte temporale und fo-basierte Messgrößen automatisch zu berechnen. Die Ergebnisse sollten in Tabellen gespeichert oder grafisch dargestellt werden. Die ausgegebenen Dateien sollten anschließend an bereits etablierte Systeme des Zentrums für vorsprachliche Entwicklung und Entwicklungsstörungen zur Weiterverarbeitung übergeben werden können.

Im Rahmen der vorliegenden Arbeit wurden basierend auf Vorgaben eines Projektes aus der klinischen Praxis zwei PRAAT-Skripte zur halbautomatischen Annotation beziehungsweise automatischen Extraktion und Analyse akustischer Kenngrößen erarbeitet, implementiert und getestet. Die beiden Skripte werden üblicherweise nacheinander verwendet, können jedoch auch getrennt voneinander eingesetzt werden. Die Vorteile und Testergebnisse werden auch in Bezug auf existierende Systeme diskutiert.

Im Laboralltag des ZVES konnte in Zusammenarbeit mit den Forschern des Projektes „Analyse von Komfortlauten (babbling) bei Säuglingen zum Zeitpunkt

des Neugeborenen-Hörscreening-2“ (Ethikvotumnummer 308/17me), die Funktionsfähigkeit des Skriptes erfolgreich getestet und die Eignung durch weitere Anwender bestätigt werden (z.B. Blum, 2020).

#### 4.1.1 Vorteile der skriptbasierten Voranalyse

Um aus den Originalaufnahmen, die unter klinischen Alltagsbedingungen erstellt werden, die für die Vorsprachliche Diagnostik relevanten Signale und Signaleigenschaften ermitteln zu können, sind umfangreiche Vorarbeiten und Voranalysen erforderlich. Dazu zählen z.B. die Editierung und Archivierung der Rohdaten, inklusive datenschutztechnisch notwendiger Pseudonymisierungs- und Anonymisierungsprozeduren sowie die Identifikation und Speicherung von Einzelvokalisationen aus den Rohdaten. Die korrekte Extraktion der Einzelvokalisationen ist im Rahmen der Voranalysen für die Vorsprachliche Diagnostik essenziell.

Der damit verbundene, bisher sehr zeitaufwändige Prozess der Identifikation von relevanten vom Säugling produzierten Einzellauten innerhalb der Originalaufnahmen (Langserien) wurde mithilfe des hier erarbeiteten Annotationsskriptes zur halbautomatischen Annotation deutlich effizienter und benutzerfreundlicher im Vergleich zur bisherigen rein manuellen Annotation. Bereits beim Einsatz der verschiedenen Vorabversionen der Skripte im Laboralltag des ZVES konnte eine deutliche Steigerung in der Verarbeitungsgeschwindigkeit der Rohdaten beziehungsweise Originalaufnahmen festgestellt werden. Wie in Kapitel 3.2 dargestellt, wurde selbst bei vergleichsweise schlechtem Signal-Rausch-Abstand die Mehrzahl der Lautgrenzen korrekt erkannt, also Einzellaute und inspiratorische Pausen korrekt identifiziert (Kapitel 4.1.2.2).

Werden die zu annotierenden Originaldateien in einen gemeinsamen Ordner abgelegt, müssen die Signale nun nicht mehr einzeln und manuell in PRAAT eingeladen werden. Das Skript bietet die Möglichkeit, unter Angabe eines Dateipfades im Eingabefenster, die Dateien automatisch in PRAAT einzuladen. In nachfolgenden Eingabefenstern können die Nutzer Parameter wie Intensitätsschwellenwert oder Suchbereich für die fo-Analyse abändern, ohne am

Skript selbst Änderungen vornehmen zu müssen. Die Möglichkeit, diese verschiedenen Parameter in Eingabefenstern abzuändern, gewährleistet eine einfache Modifikation dieser, ohne jedoch in puncto Benutzerfreundlichkeit Einbußen hinnehmen zu müssen. Dies ist für klinische Fragestellungen sehr wichtig.

Selbst wenn nach der Annotationsroutine Lautgrenzen oder Pausen manuell verschoben, entfernt oder hinzugefügt werden mussten, geben die automatisch generierten Annotationsvorschläge zumindest Anhaltspunkte, an welcher Stelle die Cursor gesetzt werden müssen. Gerade bei Signalen mit geringem Signal-Rausch-Abstand, bei denen schon bei einer rein händischen Annotation Lautgrenzen schwerlich aufzufinden sind, erwiesen sich die Vorschläge nach Angaben der Tester im laufenden Analysebetrieb als inhaltlich hilfreich und darüber hinaus zeitsparend.

Für Aufgabenstellungen, die die Reaktion eines Säuglings auf die vokale Kommunikation mit einer Kontaktperson untersuchen sollen (z.B. Imitationstest bei Säuglingen mit LKGS-Spalten oder Hörbeeinträchtigungen), müssen zeitliche Abschnitte, die zwischen den kindlichen Lautabschnitten (vokale Imitationsvorgabe) liegen, auch ausgewertet werden. Das war bisher nicht Bestandteil der Routine im ZVES-Labor. Das neue Skript ermöglicht dies durch eine variable intensitäts- und fo-basierte Detektion relevanter non-verbaler/verbaler Signale sowie die Auswertung ihrer temporalen Eigenschaften. Hier empfiehlt sich der Einsatz einer eigenen, vom Nutzer festgelegten, Nomenklatur für die kind- und elterngenerierten Passagen, um diese eindeutig von anderen Hintergrundgeräuschen abzugrenzen.

Auch das hier erarbeitete Skript zur Extraktion und Analyse der Einzellaute aus Langserien bringt mehrere Vorteile mit sich. Einzellaute müssen, im Vergleich zum bisherigen Vorgehen, nun nicht mehr manuell segmentiert und abgespeichert werden – eine enorme Zeitersparnis, die im Hinblick auf das Folgeprojekt an einer sehr großen Stichprobe gefordert war. Diese Vorgänge erfolgen direkt auf Basis der Annotationen. Da die Rohdaten so unangetastet bleiben, kann zu jedem Zeitpunkt eine beliebige Neuaufteilung oder Korrektur der



Annotation und erneute Analyse erfolgen, sollte dies notwendig werden oder inhaltlich von Interesse sein. Damit wurde die geforderte, deutlich schnellere und höhere Reproduzierbarkeit der Ergebnisse erreicht. Die analysierten Einzellaute können jetzt zusätzlich in ihrem akustischen Kontext betrachtet werden, stehen jedoch auch weiterhin noch als Einzellaute zur Verfügung, die anschließend an externe Systeme weitergegeben werden können. Auch bei diesem Skript müssen Laute vom Nutzer nicht mehr einzeln in PRAAT eingeladen werden. Es ist nun auch möglich, per Stapelverarbeitung mehrere Langserien vollautomatisch ohne weiteres Eingreifen zu analysieren.

Außerdem ermöglicht die eingebaute optional anwählbare Ausgabe temporaler Kenngrößen spezifische Rhythmusanalysen ergänzend zu den fo-Analysen. Bei der Ermittlung der temporalen Kenngrößen, wie beispielsweise Lautlänge, Melodielänge, Silben- und Wortlänge oder Pausenlänge, werden die Werte direkt vom Zeitsignal ermittelt und sind damit wesentlich genauer als jene, die basierend auf grafischen Darstellungen wie Spektren oder den fo-Werten der Melodiekurven, ermittelt werden (vgl. Kapitel 3.2.2).

Die beiden Skripte haben sich im Laboralltag des ZVES für ihren jeweils vorgesehenen Zweck als geeignet erwiesen. Der Nutzer muss jedoch eine Mindestkenntnis des Lautrepertoires der vokalen Entwicklung im ersten Lebensjahr mitbringen, um eine qualifizierte Unterscheidung der akustischen Phänomene bei einer manuellen Nachkorrektur vornehmen zu können. Es konnte ein Beitrag zur Optimierung der Vorsprachlichen Diagnostik geleistet sowie eine wesentliche methodische Voraussetzung für das hier bearbeitete und nachfolgende Forschungsprojekte geschaffen werden.

#### 4.1.2 Bewertung der Ergebnisse der automatischen Annotation

Wie in Kapitel 3.2 dargestellt, stellen spezifische Signaleigenschaften der Säuglingslaute sowie Aufnahmebedingungen im klinischen Alltag das automatische Annotationssystem vor besondere Herausforderungen. Die im Rahmen der Skripttestung hier beobachteten potenziellen Artefaktquellen sollen nachfolgend diskutiert werden.

#### 4.1.2.1 Segmentierungspausen

Die meisten Publikationen aus dem ingenieurstechnischen Bereich zur automatischen Säuglingslaut- insbesondere Schreilautanalyse, mit Ausnahme von Abou-Abbas et al. (2017), ignorieren den gut belegten Fakt, dass innerlautliche Segmentierungen bei Babylauten regelhaft vorkommen (Birr, 2009; Koopmans-van Beinum & van der Stelt, 1986; Ostermann, 2010; Robb et al., 2019; Wermke & Mende, 2009; Wermke & Mende, 2011/2012).

Das hier erarbeitete Skript detektiert Einzelvokalisationen aus den Originaltonsequenzen (Langserien) unter Verwendung der Intensität (s. Kapitel 3.1.2). Dies stellt die einfachste Möglichkeit dar, Pausen und stimmlose Elemente zu erkennen. Die bisher bekannten automatischen Babylautsysteme verwenden ebenfalls die Intensität, meist zusätzlich in Kombination mit weiteren Verfahren (Zero-Crossing Rate, Mel frequency cepstrum coefficient, empirical mode decomposition, Energie, fo) (vgl. Abou-Abbas et al., 2017). Bei vielen Vokalisationen scheint dieses einfache Verfahren bereits sehr gut zu funktionieren. Das konnte auch hier bei den Testsignalen aus verschiedenen Alters- und Patientengruppen (siehe Kapitel 3.2.1) sowie bei der Auswertung der Projektdaten demonstriert werden. In Zukunft sollten die Segmentierungspausen noch eine eigene Beschriftung innerhalb der Annotation erhalten, um auch die temporalen Kenngrößen dieser für eine Analyse verfügbar zu machen. Der Autor der vorliegenden Arbeit hat diese Erweiterung bereits im Rahmen des Kooperationsprojektes zur, in Kapitel 2.1 beschriebenen, Studie bei Säuglingen mit LKGS des ZVES in Kooperation mit dem Children's Hospital der University of Iowa begonnen und getestet (Conrad et al., 2020).

Trotz ihrer prinzipiell guten Eignung ergab sich eine potenzielle Fehlerquelle bei alleiniger Verwendung des Intensitätsmerkmals allerdings daraus, dass ein großer Teil gesunder Säuglinge regelhaft beim Weinen in den ersten drei Monaten sehr starke respiratorische Aktivitäten zeigen. Dadurch kommt es häufig zu sogenannten inspiratorischen (ingressiven) Phonationen (Wermke, Haschemi et al., 2018). Diese Lautbildung während der Inspiration und der stimmlosen Phase bis zum Beginn der egressiven Vokalisation kann zu einer fehlerhaften

Detektion segmentierter Vokalisationen bei der Sequenzanalyse führen, indem ingressive Laute fälschlich als egressiv detektiert wurden. Um diese Fehler zu vermeiden, sind bisher noch auditive Nachkontrollen notwendig. Ob komplexere Verfahren, wie das von Abou-Abbas et al. vorgeschlagene (Abou-Abbas et al., 2017), diesbezüglich leistungsfähiger sind, ist bisher nicht untersucht.

In Zusammenarbeit mit Ingenieuren und Signalanalysefachleuten sollte in einem nächsten Schritt auch die schwierigere Aufgabe gelöst werden, ein Verfahren zu entwickeln, das intensitätsarme innerlautliche Bereiche von ebenfalls intensitätsarmen Zwischenlautpausen unterscheiden kann. Das hier erarbeitete Skript arbeitet formal korrekt; inhaltliche Unterscheidungen „echter“ und „fehlerhaft“ segmentierter innerlautlicher Pausen müssen allerdings im Moment noch vom Nutzer im Nachgang getroffen und entsprechende Cursor gegebenenfalls gelöscht oder korrigiert werden (siehe Kapitel 3.2.1). Dies ist insbesondere im Zusammenhang mit der Vorsprachlichen Diagnostik bei Säuglingen mit Lippen- Kiefer- Gaumen- und Segelspalten wichtig, da diese Kinder präoperativ innerlautliche Segmentierungen deutlich häufiger zeigen als Nichtspaltträger (Ostermann, 2010; Wermke et al., 2011). Bei Nichtschreibvokalisationen in späteren Phasen der vokalen Entwicklung werden solche Segmentierungen häufig zu Platzhaltern für Konsonanten oder trennen Silben innerhalb einer Phrase beim Babbeln voneinander (Kapitel 3.2.1.4). Das gleiche Segmentierungsproblem ist damit auch bei Untersuchungen spontansprachlicher Aufnahmen älterer Kinder relevant. Bisher wurde dies allerdings noch von keiner Forschungsgruppe automatisiert gelöst.

PRAAT bietet nativ keine einfache Möglichkeit, dieser Problematik beizukommen. Hier können zukünftig möglicherweise korrekt trainierte neuronale Netzwerke dazu beitragen auch diese akustischen Ereignisse automatisiert zu erkennen.

#### 4.1.2.2 Einfluss des Signal-Rausch-Abstandes auf das Annotationsergebnis

Im klinischen Bereich sind die kindlichen Lautaufnahmen häufig durch Hintergrundgespräche und -geräusche überlagert. In der Regel sind Aufnahmen in Schallkabinen nicht möglich und schallgedämpfte Räume nicht vorhanden oder organisatorisch-logistisch im klinischen Routinebetrieb nicht nutzbar.

Die verschiedenen Hintergrundgeräusche in den Originallangserien beeinträchtigen teilweise die Annotationsvorschläge durch das Skript und markiert diese dann fälschlicherweise als relevante Passagen. Die Bandbreite solch störender Hintergrundgeräusche reicht von der Kommunikation des medizinischen Personals untereinander sowie mit den Eltern über die Eltern-Kind-Kommunikation bis zu Hintergrundgeräuschen durch medizinische Geräte oder andere Prozeduren im Therapie- und Pflegebetrieb.

Bei Aufnahmen niedriger Signalgüte (niedriger Signal-Rausch-Abstand) kommt es häufiger zu einer fehlerhaften oder fehlenden Markierung. Beispielsweise mussten hier bei Testsignal D, (Kapitel 3.2.1.1) aufgrund des niedrigen Signal-Rausch-Abstandes (vgl. auch Abbildung 31) im vorderen Teil gleich mehrere Lautgrenzen hinzugefügt werden, um die einzelnen Wörter korrekt in das PRAAT-Tier zu annotieren.

Die relevanten kindlichen Vokalisationen setzen sich bei niedrigem Signal-Rausch-Abstand nicht deutlich von den Hintergrundgeräuschen ab, was dann methodenbedingt zu Fehlern in der Annotation führt. Die Analyse erfolgt immer über die komplette aktuelle akustische Sequenz. Durch Ändern des Intensitätsschwellenwertes im Benutzermenü kann diesem Problem zum Teil entgegengewirkt werden: der Schwellenwert wird zur Identifikation der „lauten“ sowie „lautlosen“ Bereiche dann so gewählt, dass ein Großteil der Einzellaute innerhalb einer Langserie über eine Intensitätsdifferenz des eingestellten Betrages verfügt. Sind die Vokalisationen innerhalb einer solchen Sequenz stark unterschiedlich ausgesteuert oder natürlicherweise durch starke Intensitätsschwankungen charakterisiert, kommt es in der Folge wieder zu mehr fehlerhaften Vorschlägen durch das Skript. Diese müssen dann bisher noch im

weiteren Verlauf der Annotation durch den Benutzer manuell korrigiert werden, um eine zuverlässige Datenanalyse zu sichern. Testsignal B (siehe Kapitel 3.2.1.1) enthielt z.B. starke Intensitätsschwankungen (vgl. Zeitbereichsdarstellung) sowie Hintergrundgeräusche. Dadurch wurden alle Signalanfänge mit Standardeinstellungen zunächst inkorrekt (zu spät) annotiert. Testsignal C enthielt auch Hintergrundgeräusche. Trotzdem wurde die Segmentierungspause im zweiten „e“-Laut (Abbildung 29) korrekt annotiert.

Die eben genannte Problematik verschärft sich bei der Aufnahme und Analyse von Komfortvokalisationen (non-cry vocalisations). Im Unterschied zum Weinen, das nur bei sehr jungen Säuglingen im Rahmen der Vorsprachlichen Diagnostik am ZVES analysiert wird, zeichnen sich die Komfortvokalisationen (Gurren, Babbeln) ab dem 3. Lebensmonat dadurch aus, dass sie prinzipiell deutlich leiser sind als Säuglingsschreie; der Signal-Rausch-Abstand ist hier dadurch eine noch kritischere Einflussgröße im Gegensatz zur Analyse von Säuglingsweinen. Ein kritischer Punkt bei der Aufzeichnung der Komfortvokalisationen im klinischen oder auch häuslichen Kontext ist die korrekte, individualisierte Aussteuerung einzelner, spontan erzeugter Vokalisationspassagen, um einen möglichst hohen Signal-Rausch-Abstand bei vergleichsweise leisen Lauten zu erzielen (z.B. Gurrlaute).

Ein ähnliches Problem wurde bei Vokalisationen älterer Säuglinge beobachtet. Kanonische, das heißt akustisch zielsprachlich geprägte Babbellaute (die unmittelbar dem Sprechen vorausgehen), weisen charakteristische Silbenstrukturen auf, die typischerweise unterschiedlich betont sind und daher in der Intensität variieren. Es kommt oft zu markanten Intensitätslücken zwischen den Silben (z.B. Abbildung 12). Wenn Zwischensilbenbereiche oder Zwischenwortpausen bei Sätzen zu lang werden, fällt dem bisherigen System die Unterscheidung zu „echten“ Pausen noch schwerer (da der Hintergrund nicht bei 0dB SPL liegt). Die Einführung eines kombinierten Intensitäts-Zeitkriteriums könnte hier für eine eindeutigere Unterscheidung sorgen, erfordert jedoch eine individuelle Anpassung je nach Vokalisationstyp und Reife der oro-pharyngealen Koordination, die die Schnelligkeit der artikulatorischen Bewegungen bestimmt. Hier liegt ein weiterer notwendiger Forschungs- und Entwicklungsschwerpunkt.

#### 4.1.3 Vergleich mit anderen Analyseansätzen und Zukunftsaussichten

Obgleich der Einsatz der hier erarbeiteten Programmroutinen schon zu einer erheblichen Steigerung der Geschwindigkeit, sowohl bei der Annotation als auch Analyse vorsprachlicher Lautäußerungen, geführt hat, bleibt der Ansatz letztendlich ein halbautomatischer, der nach wie vor eine manuelle Prüfung und Sichtung durch einen (menschlichen) Benutzer voraussetzt. Die hohe Zuverlässigkeit dieses Vorgehens, die für klinische Fragestellungen unumgänglich ist, wurde bisher von keinem automatischen Ansatz erreicht und ist Bestandteil der laufenden Forschung am ZVES.

Bisherige Versuche die Babylautanalyse zu automatisieren, wie in Kapitel 1.1 beschrieben, berichteten vereinzelt vielversprechende Ansätze. Da die Systeme jedoch bisher ohne Berücksichtigung der inzwischen erforschten medizinisch-biologischen Hintergrundinformationen und ohne enge Zusammenarbeit mit Ärzten trainiert wurden, bleiben sie gegenwärtig noch reine technische Algorithmenproben. Keines der publizierten Verfahren hat sich, soweit den Publikationen zu entnehmen ist, einer geeigneten klinischen Testung gestellt; sie sind daher für einen klinischen Einsatz (noch) unbrauchbar.

Neuronale Netzwerke und die Kombination dieser mit stochastischen Suchmethoden bieten auf dem Feld der Spracherkennung bereits vielversprechende Resultate (Graves, Fernández, Gomez, & Schmidhuber, 2006). Sie könnten in Zukunft auch in der Verarbeitung frühkindlicher Laute eine zentrale Rolle einnehmen. Allerdings gilt dies nur unter der Voraussetzung, dass zunächst eine relevante, zuverlässige, alters- und sprachspezifische Referenzwertedatenbank erstellt wird, die es ermöglicht, Systeme geeignet anzulernen. Dazu sind Forschungen, wie sie am ZVES durchgeführt werden, unerlässlich. Die bisherigen weltweiten Forschungsaktivitäten zentrieren sich faktisch bisher, bedingt durch unterschiedliche Fachdisziplinen, auf zwei separate Felder. Zum einen werden anspruchsvolle, komplexe Algorithmen für automatische Analysen an oft medizinisch unzureichend differenzierten Testsignalen entwickelt bzw. getestet und in technischen Journalen für

Fachkollegen publiziert. Zum anderen werden geeignete Referenzwerte akribisch, unter hohem Personal- und Zeitaufwand, sowohl an gesunden als auch an auffälligen Säuglingen interaktiv ermittelt. Das ZVES versucht, beide Strategien zu verbinden. Die vorliegende Arbeit hat dazu einen Beitrag geleistet.

Ein Vorteil der Vielzahl der bisherigen Ansätze besteht zumindest darin, dass potenziell geeignete Messgrößen, insbesondere die Grundfrequenz, die Kontur der Grundfrequenzkurve (Melodie) und die Analyse temporaler Eigenschaften, bei der Suche nach Frühindikatoren für neurophysiologische Entwicklungsstörungen herausgearbeitet wurden (vgl. Übersicht in Barr et al. 2000, S. 137–156; LaGasse, Neal, & Lester, 2005). Gegenstand aktueller Forschungen sind daher insbesondere die automatische Analyse dieser akustischen Merkmale. Dabei besteht das Ziel vor allem darin, potenzielle Frühindikatoren und Marker des aktuellen Entwicklungsstandes zu bestimmen. Nur eine frühzeitige und im individuellen Fall schnelle Identifikation bei sich ankündigenden Entwicklungs- beziehungsweise Sprachentwicklungsstörungen von Risikokindern ist klinisch als nicht-invasive Zusatzinformation relevant.

Eine der größten Herausforderungen, bevor eine klinischen Maßstäben genügende verlässliche Analyse frühkindlicher Lautäußerungen bei Risikokindern erfolgen kann, besteht nach wie vor zunächst in der zuverlässigen Analyse des Variationsbereiches der akustischen Messgrößen bei unauffälligen Säuglingen und der Identifikation relevanter Einflussfaktoren. Ohne solche Referenzwerte bleibt jedes interaktive oder automatische Analyseprogramm diagnostisch unbrauchbar. Jedoch hilft die durch geeignete Programme mögliche Effizienzsteigerung auf dem Weg dahin, dieses Ziel schneller zu erreichen.

#### 4.1.4 Ermittlung der Frequenzeigenschaften mit dem Skript zur automatischen Analyse annotierter Langserien

Bei der Gegenüberstellung der Daten aus den Analyseergebnissen der Testsignale, die mit dem Skript zur Analyse der annotierten Langserien (Kapitel 3.1.1) erstellt wurden, konnten keine Unterschiede zu den fo- und Intensitätsdaten festgestellt werden, die zuvor manuell mit dem von Prochnow erstellten und routinemäßig am ZVES verwendeten Skript erzeugt wurden

(Prochnow, 2013). Dadurch, dass die Analyseergebnisse im Vergleich zur manuellen Variante jedoch zuerst in einer programmeigenen Tabelle gespeichert und dann gesammelt gespeichert werden und nicht zeilenweise aus dem Info-Window in eine Textdatei geschrieben werden, konnte die Geschwindigkeit deutlich erhöht werden. Den Flaschenhals stellt hier vermutlich die Kommunikation des Skriptes mit der Dateiverwaltung des jeweiligen Betriebssystems dar. Für die Projektdaten (vgl. Kapitel 2.1) waren nur temporale Eigenschaften gefordert, so dass nur diese im Kapitel 3.3.2 berichtet wurden.

## 4.2 Temporale Eigenschaften im Methodenvergleich

Die Vergleiche der Messungen der Vokalisationslängen ( $L_{ges}$  [s]) an den hier verwendeten Testsignalen lieferten einen mittleren Unterschied in der Zeitdauer der Vokalisationen von 21 ms (Median) zwischen der Messung an den Melodiekurven und am Zeitsignal, basierend auf dem hier erstellten Skript. Dabei waren die am Zeitsignal ermittelten Werte (SAM-Methode) mehrheitlich länger als die an den Melodiekurven (MM-Methode) ermittelten. Der Unterschied erscheint relativ klein, kann aber bei bestimmten Fragestellungen (z.B. Silbendauer, Plateaulängen von Melodie-Intervallstrukturen) inhaltlich relevant sein. Nur vier Vokalisationen der Testsignale zeigten eine längere Gesamtdauer bei Messung an den Melodien. Bei den Projektdatenanalysen bestätigte sich der Befund bezüglich der Zeitdifferenzen bei Verwendung beider Methoden, insbesondere bei den segmentierten Vokalisationen:

Über alle Melodiestrukturkategorien hinweg lieferte die semi-automatische Methode (SAM) im Mittel geringfügig längere Zeitwerte als die an den Melodiekurven vermessenen (vergleiche Kapitel 3.3.2).

Bezüglich der Pausenlängen in segmentierten Vokalisationen zeigte sich dagegen, dass die Pausenlängen in der manuellen Methode meist etwas länger waren (vergleiche Kapitel 3.3.2.4 bis 3.3.2.7). Beide Befunde basieren auf demselben Effekt der präziseren Messung am Zeitsignal mit mehr Messpunkten des Zeitsignals an den Rändern der phonierten Strukturen.



Zusätzlich wurden die Ergebnisse der Inspirationslängen in den Tabellen mit berichtet, da sie automatisch bei der Skriptverwendung mit ausgegeben werden. Da es kaum Daten dafür in der Fachliteratur gibt, können sie für andere Forscher nützlich sein.

Die Mittelwerte der Inspirationslängen lagen bei 0,367s (SD=  $\pm$  0,252s) (Postphonation). Die hier ermittelten Mittelwerte und Mediane für die Präphonationsinspiration liegen bei 0,277s (SD =  $\pm$  0,194s) und 0,248s und damit geringfügig (78ms bzw. 57ms) über den von Sereschk ermittelten Werten bei Neugeborenen (MW: 0,199s (SD=  $\pm$  0,055s) Med: 0,191) (Sereschk, 2019, pp. 28–29). Danach liegen sie in der gleichen Größenordnung, obwohl die Säuglinge im Projekt hier bereits zweieinhalb Monate alt waren.

Zusammenfassend hat die vorliegende Arbeit gezeigt, dass bei der Ermittlung temporaler Messgrößen und Charakteristika in der Vorsprachlichen Diagnostik die Messung am Zeitsignal präferiert werden sollte. Das hier erarbeitete Skript liefert dazu ein geeignetes Handwerkszeug. Selbstverständlich liefert die manuelle Cursorsetzung, sofern die automatisch ermittelten Positionen manuell korrigiert werden müssen, ebenfalls eine gewisse Messvariabilität. Diese ist jedoch deutlich geringer und liegt bei ca. 3ms (Sereschk, 2019).

### 4.3 Projektdatenanalyse

Bei den im Rahmen des Projektes (Kapitel 2.1) hier analysierten Säuglingslauten zeigte sich bei Säuglingen mit LKGS-Spaltbildungen ein deutlich höherer Anteil von Vokalisationen mit segmentierten Melodien im Vergleich zur Kontrollgruppe. Dieser Befund steht in Übereinstimmung mit vorhergehenden Untersuchungen, in denen ein hoher Anteil segmentierter Melodien als charakteristisch für Vokalisationen für Säuglinge mit LKGS- Spaltbildungen postuliert wurde (Birr, 2009; Ostermann, 2010; Wermke et al., 2011). Interessant war auch der Befund, dass sich bei generell höherem Segmentierungsanteil innerhalb der LKGS-Gruppe die Probanden untereinander unterschieden. Die vorläufige Auswertung im Rahmen des Gesamtprojektes bezüglich des Segmentierungsverhaltens in Kombination mit den MRT-Befunden postulierte eine Korrelation zwischen dem Segmentierungsanteil und dem Anteil weißer Substanz in den MRT-Befunden. („As a group, unaffected infants had a mean percentage of 12.6% cry utterances with melody breaking (laryngeal constriction), while for infants with iCL/P, over 60% of cry utterances were interrupted. [...] Overall, the two infants with iCL/P had a pattern of consistently lower volume for cerebral and cerebellar gray matter, with a higher volume of cerebral unmyelinated white matter [...]” (Conrad et al., 2020)). Damit hat das durchgeführte Pilotprojekt erstmalig Argumente dafür vorgelegt, dass es einen potenziellen Zusammenhang zwischen der Hirnreifung bzw. strukturellen Hirnorganisation und den Rhythmus-eigenschaften der Säuglingslaute geben könnte. Grundlage für dieses Postulat waren auch die hier erarbeiteten Befunde und Daten. Um dieses Postulat zu prüfen, ist bereits eine Folgestudie mit großer Stichprobe beantragt, in der die hier getestete SAM-Methode Teil des Arbeits- und Analyseprogramms ist.

Die vorliegende Arbeit hatte das Ziel, die temporalen Eigenschaften der segmentierten und nichtsegmentierten Vokalisationen der Projektprobanden im Rahmen einer Kooperation zuverlässig mit Hilfe einer neuen, semi-automatischen Methode zu ermitteln. Alle Messwerte sind ausführlich berichtet, sodass sie für Folgeprojekte und zu Referenzzwecken für andere Forschungsgruppen zur Verfügung stehen. Vergleichswerte liegen mit

Ausnahme der Arbeit von Birr (2009) in der Literatur bisher diesbezüglich nicht vor. Birr (2009) hat 14 Säuglinge mit unterschiedlichem Spalttyp im Längsschnitt über die ersten 3 Monate untersucht und einer altersadäquaten Gruppe von Säuglingen ohne Spaltbildung gegenübergestellt. In der 5. - 8. Lebenswoche, also dem Alter, das dem der hier untersuchten Säuglinge entspricht, berichtet sie einen MCI-Medianwert von 0,46 in der LKGS-Gruppe und 0,58 in der Kontrollgruppe. Das bedeute, dass die Säuglinge der Kontrollgruppe prozentual mehr komplexe Melodiestructuren in ihren Weinauten produzierten als die Säuglinge der LKGS-Gruppe. Dieser Befund weicht von den hier gefundenen Werten deutlich ab. Die Abweichungen betreffen insbesondere die LKGS-Gruppe, bei der in gleichem Alter ein deutlich höherer MCI von 0,83 gefunden wurde. Eine mögliche Erklärung für diese Unterschiede könnte in der kleinen Stichprobe hier bei gleichzeitig hoher interindividueller Variabilität liegen. Der MCI der Kontrollgruppe hier (0,47) entspricht dem von Birr (2009) berichteten.

Birr (2009) hat gezeigt, dass alle Melodiestructurkategorien, die man nach dem vorsprachlichen Entwicklungsprogramm erwarten würde (Wermke, 2002; Wermke & Mende, 2011/2012), auch bei Säuglingen mit orofazialen Spaltbildungen zu beobachten sind. Das hat auch Ostermann (2010) gefunden, allerdings aufgezeigt, dass die höher komplexen Kombinationsmuster bei Säuglingen mit orofazialen Spaltbildungen in den ersten Lebensmonaten seltener vorzukommen scheinen. Das wurde hier nicht gefunden (Kapitel 3.3.1). Birr (ibid.) beschreibt, dass das Verhältnis von nichtsegmentierten zu segmentierten Melodien in der Kontrollgruppe über die ersten drei Lebensmonate bei ungefähr 1 liegt, also etwa gleich viele segmentierte wie nicht segmentierte Strukturen erzeugt wurden. Auch Lind (2007) kommt zu einem ähnlichen Ergebnis. Die Kinder mit LKGS-Spaltbildungen zeigten dagegen ein Verhältnis nichtsegmentierter/segmentierter Laute von 0,65 bis 0,87 im Alter von 5 bis 12 Lebenswochen (Birr, 2009). Den Hauptanteil der segmentierten Kategorien bildeten die 1S-Strukturen, wie auch in der hier durchgeführten Analyse. Im Unterschied zu den Kontrollprobanden in der hier vorliegenden Arbeit, zeigten alle Probanden der LKGS-Gruppe einen deutlich höheren Anteil aller segmentierten Kategorien (1S, 2S, 3S, KS). Birr (ibid.) erklärt diesen erhöhten

1S-Anteil indem sie postuliert, dass diese Strukturategorie eine Art „Ersatzstrategie“ für fehlende Mehrfachbogenstrukturen darstellen könnte, da für diese Säuglinge segmentierte Strukturen leichter zu produzieren sein könnten als doppel- und mehrböjige Melodien. Die Pilotstudie des Forschungsprojektes, in das die vorliegende Arbeit eingebettet ist, und weitere Veröffentlichungen deuten dagegen eher auf hirnmorphologische Besonderheiten der Säuglinge mit LKGS-Spaltbildungen hin, die u.a. erstmals einen verringerten Myelenisierungsgrad und eine geringeres Volumen des Cerebellums bei den LKGS-Probanden postulieren (Conrad et al., 2010; Nopoulos et al., 2017; Peg Nopoulos, Douglas R. Langbehn, John Canady, Magnotta, & Lynn Richman, 2007; Yang, McPherson, Shu, Xie, & Xiang, 2012). Aufgrund der hohen interindividuellen Variabilität in der frühen Sprech- und Sprachentwicklung bei medizinisch unauffälligen Säuglingen und Säuglingen mit orofazialen Spaltbildungen müssen nachfolgende Studien analysieren inwiefern dieses Postulat einer möglicherweise nur transitorischen Entwicklungsbesonderheit tatsächlich generalisierbar ist. Die sehr geringe Stichprobenzahl der durchgeführten Pilotstudie erfordert eine äußerst zurückhaltende Interpretation der korrelierten Befunde zwischen Hirnstrukturuntersuchungen und vorsprachlichen Lautanalysen. Allerdings gibt es bei älteren Patienten Studien, die das Postulat stützen (Nopoulos et al., 2002; Shriver, Canady, Richman, Andreasen, & Nopoulos, 2006).

Die einfachsten hier untersuchten Strukturen sind Vokalisationen, deren Melodie aus einem einfachen einzelnen Melodiebogen bestehen. Bei diesen einfachen Strukturen würde man die geringsten Unterschiede zwischen beiden Probandengruppen erwarten. Die hier durchgeführte Analyse zeigte jedoch, dass bereits deutliche Unterschiede in den Verteilungseigenschaften der Gesamtlänge dieser einfachen 1B-Vokalisationen zwischen beiden Probandengruppen, zumindest auf Sigallevel, bestehen. Der Medianwertunterschied zwischen beiden 1B-Probandengruppen betrug 336ms (SAM-Methode). Die Probanden der LKGS-Gruppe produzierten deutlich längere 1B-Vokalisationen. Der Mittelwert der 1B-Laute betrug bei der KG-Gruppe  $0,61 \pm 0,3s$ , bei der LKGS-Gruppe  $0,94 \pm 0,28s$ . Alle Analysen sind allerdings nur auf Signallevel durchgeführt, da es in der LKGS-Gruppe nur insgesamt 11 einböjige Vokalisationen gab (Proband 1

n=4; Proband 2 n=0; Proband 3 n=6; Proband 4 n=1). Die Interpretation der Verlangsamung der Phonation muss dementsprechend sehr zurückhaltend sein. Aus Untersuchungen bei erwachsenen Probanden ist bekannt, dass das Cerebellum bei Diskriminationsaufgaben in Hörtests bei Patienten mit diagnostizierter cerebellarer Degeneration im Vergleich zu Kontrollprobanden deutlich langsamer ist. Betroffene Patienten zeigten eine 285ms langsamere Diskrimination von Frequenzunterschieden als die Kontrollen (allerdings nicht statistisch signifikant (Parsons, Petacchi, Schmahmann, & Bower, 2009). Das ist zwar nicht direkt vergleichbar mit den hier erhobenen Befunden, bestätigt allerdings indirekt die Beteiligung des Cerebellums an auditiven Diskriminierungsleistungen und eine Verzögerung der Prozesse bei eingeschränkter Leistungsfähigkeit. Interessanterweise gibt es Befunde anderer Autoren zur Rolle des Cerebellums bei den Sprachleistungen von Spaltpatienten (siehe z.B. Dirauf, 2010; Nopoulos et al., 2002; Steck-Walter, 2007).

Ein weiteres stützendes Argument für das Postulat der potenziellen Verlangsamung der Phonationsprozesse lieferte der Befund, dass sich hier generell längere Zeiten auch für eine Reihe anderer Strukturkategorien, z.B. 3B/MB, 1S, 3S, und melodische Teilelemente fanden (Kapitel 3.3.2.3, 3.3.2.4 und 3.3.2.6).

## 5 Zusammenfassung

Seit mehr als 60 Jahren beschäftigen sich verschiedene Forschergruppen weltweit mit der Vorsprachlichen Diagnostik, also der Analyse und Auswertung frühkindlicher Lautäußerungen, mit dem Ziel, eine nicht-invasive Methode zur Früherkennung von Entwicklungsstörungen des auditiv-vokalen Systems zu entwickeln, die medizinische Verfahren im Säuglingsalter ergänzen könnte. Die Erstellung einer geeigneten Referenzdatenbank des normalen Variationsbereiches akustischer Kenngrößen von Säuglingslauten, die auch dazu beitragen könnte vollautomatische Klassifizierungssysteme korrekt trainieren zu können, stellt aber nach wie vor eine der größten noch zu bewältigenden Herausforderungen in dieser Forschungsdisziplin dar. Die Komplexität des Informationsgehaltes von Säuglingslauten und die sehr hohe intra- und interindividuelle Variabilität von Lauteigenschaften hat existierende Ansätze immer wieder an ihre Grenzen gebracht und ihre klinische Nutzung bisher stark limitiert.

Um eine zuverlässige Analyse vorsprachlicher Lautäußerungen vornehmen zu können, müssen die einzelnen egressiven Vokalisationen der Säuglinge zunächst aus den im klinischen Kontext erfassten Aufnahmen identifiziert, von Hintergrundgeräuschen und ingressiven Phonationen bzw. inspiratorischen Geräuschen unterschieden und schließlich sinnvoll archiviert werden. Leider steht hierfür bisher noch kein marktreifes, klinisch relevantes automatisches System zur Verfügung, was der Durchführung von Studien mit großer Fallzahl im Wege steht.

Die im Rahmen der vorliegenden Dissertation erarbeiteten PRAAT-Skripte sollen, integriert in bestehende Analysesysteme, in einem interdisziplinären Setting einen Beitrag zur angestrebten Automatisierung der Vorsprachlichen Diagnostik leisten.

Das hier erarbeitete und getestete Skript zur semi-automatischen Annotation von Langserien kombiniert automatisches Einladen und Annotieren von im klinischen Kontext erfassten Lautäußerungen (Langserien) mit einer komfortablen Kontrolle und Anpassung der automatisch generierten Annotationsvorschläge durch einen

Benutzer. Mit dem Skript zur Analyse annotierter Langserien lassen sich anschließend beliebige gewünschte annotierte Lautabschnitte aus den annotierten Langserien extrahieren, mit denen anschließend Grundfrequenz-, Intensitäts- und temporale Analysen durchgeführt werden können. Die extrahierten Daten können im Anschluss abgespeichert werden und zur weiteren Analyse direkt an andere Systeme oder Statistiksoftware weitergegeben werden.

Die geforderte, ausführliche Dokumentation im Rahmen dieser Dissertation soll zukünftigen Anwendern die Anwendung und das Verständnis der Skripte erleichtern.

Die zuverlässige Funktion der Skripte wurde anhand von vorgegebenen Testsignalen aus dem Archiv des ZVES demonstriert. Auch in aktuellen Forschungsprojekten bei Säuglingen am UKW im Rahmen von Kooperationsprojekten konnten die Skripte bereits eingesetzt und auf ihre korrekte Funktion überprüft werden.

Dieser eher methodische Teil der Aufgabenstellung diente der Vorbereitung eines größeren Forschungsprojektes. Hier wurden Aufnahmen einer Pilotstudie des ZVES und dem Stead Family Department of Pediatrics der University of Iowa Roy J and Lucille A Carver College of Medicine (Iowa City, IA, USA) mit Hilfe der neu erarbeiteten und umfangreich getesteten Methode analysiert und verschiedene temporale Messgrößen detailliert berichtet. Die erstellten Tabellen liefern erstmalig Referenzwerte für diese Messgrößen.

Analyseergebnisse der vorliegenden Arbeit sind in Publikationen des Projektes eingeflossen. Sie haben es ermöglicht, erstmalig einen potenziellen Zusammenhang zwischen dem Segmentierungsverhaltens in der Phonation und der Hirnreifung bei Säuglingen im Alter von 2 Monaten zu postulieren. Aufgrund des Pilotstudiencharakters ist diese Hypothese nur vorläufig, ein Folgeprojekt unter Einbeziehung der hier erarbeiteten SAM-Methode aber bereits geplant.

## 6 Literaturverzeichnis

- Abou-Abbas, L., Tadj, C., & Fersaie, H. A. (2017). A fully automated approach for baby cry signal segmentation and boundary detection of expiratory and inspiratory episodes. *The Journal of the Acoustical Society of America*, 142(3), 1318.
- Barr, R. G., Hopkins, B., & Green, J. A. (Eds.) (2000). *Crying as a sign, a symptom, & a signal: Clinical, emotional and developmental aspects of infant and toddler crying. Clinics in developmental medicine: Vol. 152*. London: Mac Keith Press.
- Birr, M. (2009). *Melodiekomplexität - ein geeigneter Frühindikator potentieller Sprachentwicklungsstörungen bei Säuglingen mit orofazialen Spalten?* (Dissertation). Universität Würzburg, Würzburg.
- Blum, D. M. (2020). *Vergleichende Analyse melodischer Eigenschaften von frühen Komfortlauten bei gesunden und hochgradig hörgeschädigten Säuglingen* (Dissertation, eingereicht) Universität Würzburg, Würzburg.
- Boersma, P., & Weenink, D. (2018a). PRAAT (Version 6.0.3.7) [Computer software]. Retrieved from <http://www.praat.org/> (24.05.2018).
- Boersma, P., & Weenink, D. (2018b). PRAAT Manual Intro. Retrieved June 06, 2018, from <http://www.fon.hum.uva.nl/praat/manual/Intro.html>.
- Boersma, P., & Weenink, D. (2018c). PRAAT Scripting Manual. Retrieved May 24, 2018, from <http://www.fon.hum.uva.nl/praat/manual/Scripting.html>.
- Boersma, P., & Weenink, D. (2019). PRAAT Manual Sound: To Pitch (ac). Retrieved May 07, 2019, from [http://www.fon.hum.uva.nl/praat/manual/Sound\\_\\_To\\_Pitch\\_\\_ac\\_\\_\\_\\_.html](http://www.fon.hum.uva.nl/praat/manual/Sound__To_Pitch__ac____.html).
- Borysiak, A., Hesse, V., Wermke, P., Hain, J., Robb, M., & Wermke, K. (2017). Fundamental Frequency of Crying in Two-month-old Boys and Girls: Do Sex Hormones During Mini-puberty Mediate Differences? *Journal of Voice : Official Journal of the Voice Foundation*, 31(1), 128.e21-128.e28.



- Bosma, J. F., Truby, H. M., & Lind, J. (1965). Cry motions of the newborn infant. *Acta Paediatrica*, 54(Suppl. 135), 60–92.
- Buder, E. H., Warlaumont, A. S., & Oller, D. K. (2013). An acoustic phonetic catalog of prespeech vocalizations from a developmental perspective. In A. A. N. MacLeod & B. Peter (Eds.), *Languages and linguistics. Comprehensive perspectives on speech sound development and disorders: Pathways from linguistic theory to clinical practice* (pp. 103–134). New York: Nova Publishers.
- Byers-Heinlein, K., Burns, T. C., & Werker, J. F. (2010). The roots of bilingualism in newborns. *Psychological Science*, 21(3), 343–348.
- Clement, C. J. (2004). *Development of vocalizations in deaf and normally hearing infants. International series: Vol. 100*. Utrecht: LOT.
- Conrad, A. L., Dailey, S., Richman, L., Canady, J., Karnell, M. P., Axelson, E., & Nopoulos, P. (2010). Cerebellum Structure Differences and Relationship to Speech in Boys and Girls with Nonsyndromic Cleft of the Lip and/or Palate. *The Cleft Palate-Craniofacial Journal*, 47(5), 469–475.
- Conrad, A. L., Wermke, K., Eisenmann, M., Kuhlmann, E., Benavides, A., Kosciak, T., & Magnotta, V. (2020). Preliminary evaluation of pre-speech and neurodevelopmental measures in 7–11-week-old infants with isolated oral clefts. *Pediatric Research*, 1–6.
- Conrad, A. L., Wermke, K., Kuhlmann, E., & Eisenmann, M. (2019, June). *Language and Neurodevelopmental Measures in 7-9 week old Infants with Isolated Oral Clefts: (Posterpräsentation)*. European Cleft Palate Craniofacial Association, Utrecht.
- DeCasper, A. J., & Spence, M. J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior and Development*, 9(2), 133–150.
- Dirauf, S. (2010). *Strukturmerkmale der Melodie und des Rhythmus in vorsprachlichen Lautäußerungen* (Dissertation). Universität Würzburg, Würzburg.

- Eiband, D. *Charakterisierung vorsprachlicher Komfortvokalisationen von Kindern mit orofazialen Spaltbildungen vor und nach operativem Gaumenverschluss* (Dissertation). Universität Würzburg, Würzburg.
- Esteve-Gibert, N., & Prieto, P. (2013). Prosody signals the emergence of intentional communication in the first year of life: Evidence from Catalan-babbling infants. *Journal of Child Language*, *40*(5), 919–944.
- Fuamenya, N. A. (2011). *Quantitative analysis of subharmonic and noise phenomena in vocalizations of young infants: Comparing infants with and without orofacial clefts* (Dissertation). Universität Würzburg, Würzburg.
- Gaultier, C. (1995). Respiratory muscle function in infants. *The European Respiratory Journal*, *8*(1), 150–153.
- Golub, H. L., & Corwin, M. J. (1982). Infant cry: A clue to diagnosis. *Pediatrics*, *69*(2), 197–201.
- Graves, A., Fernández, S., Gomez, F., & Schmidhuber, J. (2006). Connectionist temporal classification. In W. Cohen (Ed.), *Proceedings of the 23rd international conference on Machine learning* (pp. 369–376). New York, NY: ACM.
- Grunwell, P. (1993). *Analysing cleft palate speech. Studies in disorders of communication*. London: Whurr.
- Hariharan, M., Sindhu, R., & Yaacob, S. (2012). Normal and hypoacoustic infant cry signal classification using time-frequency analysis and general regression neural network. *Computer Methods and Programs in Biomedicine*, *108*(2), 559–569.
- Hirst, D. (2011). The analysis by synthesis of speech melody: from data to models. *Journal of Speech Sciences*. (1(1)), 55-83.
- Höing, S. (2010). *Charakterisierung akustischer Eigenschaften des spontanen Neugeborenen schreies in den ersten drei Lebenstagen in Abhängigkeit vom Geburtsmodus* (Dissertation). Universität Würzburg, Würzburg.
- Hotz, M. M., & Gnoinski, W. M. (1979). Effects of early maxillary orthopaedics in coordination with delayed surgery for cleft lip and palate. *Journal of Maxillofacial Surgery*, *7*(3), 201–210.

- Howard, S. (Ed.) (2011). *Cleft palate speech: Assessment and intervention* (1. publ). Chichester: Wiley-Blackwell.
- KAY Pentax Medical. Computerized Speech Lab CSL™ (Version 4500) [Computer software]. USA: KAY Pentax Medical.
- Kisilevsky, B. S., Hains, S. M. J., Lee, K., Xie, X., Huang, H., Ye, H. H., . . . Wang, Z. (2003). Effects of experience on fetal voice recognition. *Psychological Science*, 14(3), 220–224.
- Koch, J., Koch, H., Grzonka, M., & Gundlach, K. K. H. (2003). Die LKGS(N)-Spalten und ihre Kodierung mit der LAHS-Nomenklatur [Facial clefts and their coding with LAHS nomenclature]. *Mund-, Kiefer- und Gesichtschirurgie : MKG*, 7(6), 339–344.
- Komposch, G. (1986). Die prächirurgische kieferorthopädische Behandlung von Säuglingen mit Lippen-Kiefer-Gaumen-Spalte. *Fortschritte der Kieferorthopädie*, 47(5), 362–369.
- Koopmans-van Beinum, F. J., & van der Stelt, J. M. (1986). Early Stages in the Development of Speech Movements. In B. Lindblom & R. Zetterström (Eds.), *Precursors of Early Speech* (pp. 37–50). London: Palgrave Macmillan UK.
- LaGasse, L. L., Neal, A. R., & Lester, B. M. (2005). Assessment of infant cry: Acoustic cry analysis and parental perception. *Mental Retardation and Developmental Disabilities Research Reviews*, 11(1), 83–93.
- Lederman, D., Zmora, E., Hauschildt, S., Stellzig-Eisenhauer, A., & Wermke, K. (2008). Classification of cries of infants with cleft-palate using parallel hidden Markov models. *Medical & Biological Engineering & Computing*, 46(10), 965–975.
- Lennes, M. (2003a). collect\_formant\_data\_from\_files.praat [Computer software]. Retrieved from <https://lennes.github.io/spect/> (27.05.2020).
- Lennes, M. (2003b). collect\_pitch\_data\_from\_files.praat [Computer software]. Retrieved from <https://lennes.github.io/spect/> (26.11.2018).
- Lind, K. (2007). *Untersuchung spektraler und melodischer Eigenschaften von Säuglingsschreien. Eine Longitudinalstudie über die ersten 16*

*Lebenswochen im Rahmen der Deutschen Sprachentwicklungsstudie*  
(Dissertation). Humboldt-Universität zu Berlin, Berlin.

- Linz, C., Schweitzer, T., Brenner, L. C., Kunz, F., Meyer-Marcotty, P., & Wermke, K. (2018). Does shape affect function? Articulatory skills in babbling of infants with deformational plagiocephaly. *Child's Nervous System : ChNS : Official Journal of the International Society for Pediatric Neurosurgery*, 34(3), 503–510.
- Mahrt, T. (2016). 'Append to file' is 10 times slower than writing a table. Retrieved April 14, 2020, from <https://praat-users.yahogroups.co.narkive.com/Vnxsbuxj/append-to-file-is-10-times-slower-than-writing-a-table>.
- 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.
- Mayer, J. (2012). syllable candidates. Retrieved May 27, 2020, from [http://praatpfanne.lingphon.net/downloads/syllable\\_candidates.txt](http://praatpfanne.lingphon.net/downloads/syllable_candidates.txt).
- Mayer, J. (2017). *Phonetische Analysen mit Praat Ein Handbuch für Ein- und Umsteiger*. Retrieved May 27, 2020, from <http://praatpfanne.lingphon.net/>.
- Mayer, J. (2018). *Phonetics on Speed: Praat Scripting Tutorial*. Retrieved May 24, 2018, from <http://praascripting.lingphon.net>.
- Mende, W., Herzel, H., & Wermke, K. (1990). Bifurcations and chaos in newborn infant cries. *Physics Letters A*, 145(8-9), 418–424.
- Nopoulos, P., Berg, S., Canady, J., Richman, L., van Demark, D., & Andreasen, N. C. (2017). Abnormal Brain Morphology in Patients with Isolated Cleft Lip, Cleft Palate, or Both: A Preliminary Analysis. *The Cleft Palate-Craniofacial Journal*, 37(5), 441–446.
- Nopoulos, P., Berg, S., VanDemark, D., Richman, L., Canady, J., & Andreasen, N. C. (2002). Cognitive dysfunction in adult males with non-syndromic clefts of the lip and/or palate. *Neuropsychologia*, 40(12), 2178–2184.

- Ostermann, T. (2010). *Die Analyse segmentierter vorsprachlicher Melodiestructuren von Säuglingen mit orofazialen Spalten im Alter von 2 - 3 Monaten* (Dissertation). Universität Würzburg, Würzburg
- Pachtner, S. (2017). *Quantitative Charakterisierung akustischer und phonetischer Eigenschaften kanonischer Babblers von gesunden Kindern mit deutscher Umgebungssprache* (Dissertation). Universität Würzburg, Würzburg.
- Parsons, L. M., Petacchi, A., Schmahmann, J. D., & Bower, J. M. (2009). Pitch discrimination in cerebellar patients: Evidence for a sensory deficit. *Brain Research, 1303*, 84–96.
- Peg Nopoulos, Douglas R. Langbehn, John Canady, Magnotta, V., & Lynn Richman (2007). Abnormal Brain Structure in Children With Isolated Clefts of the Lip or Palate. *Archives of Pediatrics & Adolescent Medicine, 161*(8), 753–758.
- Pelzer, A.-M. (2011). *Längsschnittanalyse der frühen sprachlichen Entwicklung bis zur Wortproduktion von Säuglingen mit orofazialen Spalten* (Dissertation). Universität Würzburg, Würzburg.
- Petroni, M., Malowany, A. S., Johnston, C. C., & Stevens, B. J. (1995). Classification of infant cry vocalizations using artificial neural networks (ANNs). In *ICASSP '95: [proceedings] 1995 IEEE International Conference on Acoustics, Speech, and Signal Processing, May 9-12, 1995, Detroit, Michigan* (pp. 3475–3478). New York, N.Y: Institute of Electrical and Electronics Engineers.
- Prochnow, A. (2013). *Der Erwerb melodisch-rhythmischer Grundbausteine im Rahmen der vorsprachlichen Entwicklung: Eine vergleichende Analyse der Schreie von schwedischen und deutschen Neugeborenen* (Dissertation). Universität Würzburg, Würzburg.
- Reyes-Galaviz, O. F., Cano-Ortiz, S. D., & Reyes-García, C. A. (2008). Evolutionary-Neural System to Classify Infant Cry Units for Pathologies Identification in Recently Born Babies. In A. Gelbukh (Ed.), *Seventh Mexican International Conference on Artificial Intelligence, 2008: MICAI '08* ;

- proceedings of the special session ; October 27 - 31, 2008, Atizapán de Zaragoza, Mexico* (pp. 330–335). Piscataway, NJ: IEEE.
- Robb, M. P., Sinton-White, H., & Kaipa, R. (2011). Acoustic estimates of respiration in the pain cries of newborns. *International Journal of Pediatric Otorhinolaryngology*, 75(10), 1265–1270.
- Robb, M. P., Yavarzadeh, F., Schluter, P. J., Voit, V., Shehata-Dieler, W., & Wermke, K. (2019). Laryngeal Constriction Phenomena in Infant Vocalizations. *Journal of Speech, Language, and Hearing Research*, 63(1), 49–58.
- Sahak, R., Lee, Y. K., Mansor, W., Yassin, A. I. M., & Zabidi, A. (2010). Optimized Support Vector Machine for classifying infant cries with asphyxia using Orthogonal Least Square. In M. K. Hamzah (Ed.), *2010 International Conference on Computer Applications and Industrial Electronics, (ICCAIE): 5 - 8 Dec. 2010, Kuala Lumpur, Malaysia* (pp. 692–696). Piscataway, NJ: IEEE.
- Sanfeliu, A., Ruiz-Shulcloper, J., Orozco-García, J., & Reyes-García, C. A. (Eds.) (2003). *A Study on the Recognition of Patterns of Infant Cry for the Identification of Deafness in Just Born Babies with Neural Networks: Progress in Pattern Recognition, Speech and Image Analysis*: Springer Berlin Heidelberg.
- Schönweiler, R., Schönweiler, B., Schmelzeisen, R., & Ptok, M. (1995). Sprach- und Sprechleistungen bei 417 Kindern mit Spaltbildungen. *Fortschritte der Kieferorthopädie*, 56(1), 1–6.
- Schwenzer, N., & Bacher, M. (Eds.) (2011). *Mund-Kiefer-Gesichtschirurgie: 48 Tab* (4., vollst. überarb. und erw. Aufl.). *Zahn-Mund-Kiefer-Heilkunde*. Stuttgart: Thieme.
- Sereschk, N. (2019). *Temporale Eigenschaften ingressiver und egressiver Phonationsleistungen gesunder Neugeborener: Temporal characteristics of inspiratory and expiratory phonation in healthy neonates* (Dissertation). Universität Würzburg, Würzburg.

- Shehata-Dieler, W., & Wermke, K. (2017). From crying to word production – observational study of an identical twin pair with severe hearing disorder. *Speech, Language and Hearing, 21*(2), 81–85.
- Shinya, Y., Kawai, M., Niwa, F., & Myowa-Yamakoshi, M. (2014). Preterm birth is associated with an increased fundamental frequency of spontaneous crying in human infants at term-equivalent age. *Biology Letters, 10*(8).
- Shriver, A. S., Canady, J., Richman, L., Andreasen, N. C., & Nopoulos, P. (2006). Structure and function of the superior temporal plane in adult males with cleft lip and palate: Pathologic enlargement with no relationship to childhood hearing deficits. *Journal of Child Psychology and Psychiatry, 47*(10), 994–1002.
- Steck-Walter, S. M. (2007). *Quantitative Strukturanalyse vorsprachlicher Vokalisationen von Säuglingen mit orofazialen Spalten im zweiten Lebenshalbjahr bei kieferorthopädischer Frühbehandlung* (Dissertation), Universität Würzburg, Würzburg.
- Vink, F. *Vorsprachliche Diagnostik im zweiten Lebenshalbjahr - Ein Vergleich zwischen Kindern mit isolierter Sagittalnahtsynostose vor und nach operativer Therapie, Säuglingen mit lagerungsbedingtem Plagiocephalus und Kontrollkindern* (Dissertation). Universität Würzburg, Würzburg.
- Wasz-Hockert, O., Lind, J., Partanen, T., Valanne, E., & Vuorenkoski, V. (1968). *The Infant Cry: A Spectrographic and Auditory Analysis: 29 (Clinics in Developmental Medicine (Mac Keith Press))* (1st ed.): Mac Keith Press.
- Wasz-Höckert, O., Michelsson, K., & Lind, J. (1985). Twenty-Five Years of Scandinavian Cry Research. In B. M. Lester & C. F. Zachariah Boukydis (Eds.), *Infant Crying: Theoretical and Research Perspectives* (pp. 83–104). Boston, MA: Springer US.
- Wendler, J. (Ed.) (2015). *Lehrbuch der Phoniatrie und Pädaudiologie* (5., unveränd. Aufl.). Stuttgart: Thieme.
- Wermke, K., & Mende, W. (2012). From Emotion to Notion: The Importance of Melody. In J. Decety & J. T. Cacioppo (Eds.), *The Oxford Handbook of Social*

- Neuroscience* (pp. 624–648). Oxford University Press (Original work published 2011).
- Wermke, K., & Mende, W. (2016). From Melodious Cries to Articulated Sounds: Melody at the Root of Language Acquisition. In M. C. Fonseca-Mora & M. Gant (Eds.), *Melodies, Rhythm and Cognition in Foreign Language Learning* (1st ed., pp. 24–47). Newcastle-upon-Tyne: Cambridge Scholars Publishing.
- Wermke, K. (1987). *Begründung und Nachweis der Eignung des Säuglingsschreies als Indikator für zentralnervöse Funktionsstörungen des Neugeborenen : Fallstudien unter Einsatz e. speziellen Computerverfahrens* (Dissertation). Humboldt-Universität zu Berlin, Berlin.
- Wermke, K. (2002). *Untersuchung der Melodieentwicklung im Säuglingsschrei von monozygoten Zwillingen in den ersten 5 Lebensmonaten* (Habilitation).
- 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.
- Wermke, K., Hain, J., Oehler, K., Wermke, P., & Hesse, V. (2014). Sex hormone influence on human infants' sound characteristics: Melody in spontaneous crying. *Biology Letters*, 10(5), 20140095.
- Wermke, K., Haschemi, A. A., Hesse, V., & Robb, M. P. (2018). Inspiratory Phonation in Baby Voice. *Journal of Voice : Official Journal of the Voice Foundation*, 32(2), 185–191.
- 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.
- 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.



- Wermke, K., Linz, C., Hasenberg, A., Kunz, F., Meyer-Marcotty, P., & Schweitzer, T. (2017). Six month-old infants with deformational plagiocephaly do not differ from unaffected infants with respect to vocal control. *International Journal of Pediatric Otorhinolaryngology*, *102*, 15–20.
- Wermke, K., & Mende, W. (2009). Musical elements in human infants' cries: In the beginning is the melody. *Musicae Scientiae*, *13*(2\_suppl), 151–175.
- Wermke, K., Pachtner, S., Lamm, B., Voit, V., Hain, J., Kärtner, J., & Keller, H. (2013). Acoustic properties of comfort sounds of 3-month-old Cameroonian (Nso) and German infants. *Speech, Language and Hearing*, *16*(3), 149–162.
- Wermke, K., Quast, A., & Hesse, V. (2018). From melody to words: The role of sex hormones in early language development. *Hormones and Behavior*, *104*, 206–215.
- Wermke, K., & Robb, M. P. (2010). Fundamental frequency of neonatal crying: Does body size matter? *Journal of Voice : Official Journal of the Voice Foundation*, *24*(4), 388–394.
- Wermke, K., Sereschk, N., May, V., Salinger, V., Sanchez, M. R., Shehata-Dieler, W., & Wirbelauer, J. (2019). The Vocalist in the Crib: The Flexibility of Respiratory Behaviour During Crying in Healthy Neonates. *Journal of Voice : Official Journal of the Voice Foundation*. Advance online publication.
- Wermke, K., Teiser, J., Yovsi, E., Kohlenberg, P. J., Wermke, P., Robb, M., . . . Lamm, B. (2016). Fundamental frequency variation within neonatal crying: Does ambient language matter? *Speech, Language and Hearing*, *19*(4), 211–217.
- Yang, F. F., McPherson, B., Shu, H., Xie, N., & Xiang, K. (2012). Structural Abnormalities of the Central Auditory Pathway in Infants with Nonsyndromic Cleft Lip and/or Palate. *The Cleft Palate-Craniofacial Journal*, *49*(2), 137–145.
- Zabidi, A., Khuan, L. Y., Mansor, W., Yassin, I. M., & Sahak, R. (2010). Classification of Infant Cries with Asphyxia Using Multilayer Perceptron Neural Network. In *Second International Conference on Computer Engineering and Applications (ICCEA), 2010: 19 - 21 March 2010, [Tuban,*

*South Kuta], Bali Island, Indonesia ; proceedings ; [including] the Second International Conference on Telecom Technology and Applications (ICTTA 2010), the Second International Conference on Knowledge Discovery (ICKD 2010) and the Second IACSIT International Workshop on Information and Emerging Technologies (IWIET 2010) (pp. 204–208). Piscataway, NJ: IEEE.*

## Danksagung

Ich danke besonders Prof. Dr. Kathleen Wermke, die durch ihre fachliche und persönliche Unterstützung wesentlich zum Gelingen dieser Dissertation beigetragen hat.

Peter Wermke danke ich für die technische Unterstützung und die Bereitstellung des CDAP-Systems.

Prof. Dr.-Ing. Mario Cebulla möchte ich für die fachliche Unterstützung danken.

Ich danke meinen Eltern Dr. Elisabeth Gradl-Eisenmann und Dr. August Eisenmann sowie Laura Geim für die Unterstützung in allen Lebenslagen.



## Eigene wissenschaftliche Tätigkeit

Conrad, Amy L.; Wermke, Kathleen; Kuhlmann, Emily; Eisenmann, Michael,  
Posterpräsentation: “Language and Neurodevelopmental Measures in 7-9 week  
old Infants with Isolated Oral Clefts” auf dem Kongress der *European Cleft Palate  
Craniofacial Association* vom 12.06-15.06.2019 in Utrecht (identifizier: ECPCA-  
672)

Amy L. Conrad; Kathleen Wermke; Michael Eisenmann; Emily Kuhlmann;  
Amanda Benavides; Timothy Kosciak; Vincent Magnotta: “Preliminary evaluation  
of pre-speech and neurodevelopmental measures in 7–11-week-old infants with  
isolated oral clefts” *Pediatric Research* (2020), 1-6

# Anhang

## A 1 Ergänzungen Skriptbeschreibung

### A 1.1 Eingabefenster

Nach dem initialen Kommentar (Zeilen 1-3, Codeabschnitt 1) (gekennzeichnet durch „#“), der hier kurz die Hauptfunktion des Skriptes beschreibt, beginnt der Skriptabschnitt für das Eingabefenster, gekennzeichnet durch „*form*“ und gefolgt von einem Textargument, hier „*Extract sounds, save, write Pitch and Intensity to table.*“ (Zeile 4), das nach Ausführung des Skriptes die Kopfzeile des Eingabefensters bildet (siehe Abbildung 18).

Die folgenden Befehle erzeugen, je nach gewünschter Funktion, unterschiedliche Eingabefelder, Texte oder Schaltflächen im Eingabefenster:

Der Befehl „*comment*“, gefolgt von einem Text, resultiert in einer Textzeile im Eingabefenster, hier zum Beispiel „*Directory of sound files*“, die als Überschrift für die Eingabefelder dient (Zeile 5).

Der Befehl „*text*“, gefolgt von einer (String-) Variablen und einem initialen Wert, schreibt den Wert im von ihm erzeugten Eingabefeld in eine String-Variable (siehe Abbildung 18). Der Name der Variablen wird jedoch im Eingabefenster nicht dargestellt (Zeile 8).

Der Befehl „*sentence*“ verhält sich analog zum Befehl „*text*“ mit dem Unterschied, dass der Name der Stringvariablen auch im Eingabefenster erscheint. Hier zum Beispiel „*output file extension*“ (Zeile 17).

Die Werte, die letztendlich in einer Stringvariablen gespeichert werden, können und sollen vom Benutzer des Skriptes durch Texteingaben bzw. Klicken in die Felder der Eingabeform noch angepasst werden.

Der Befehl „*optionmenu*“ gefolgt von einer Variablen erzeugt im Eingabefenster ein Dropdownmenü, aus dem verschiedene Optionen ausgewählt werden können. Die verfügbaren Optionen werden, passend zum Anwendungszweck, in den folgenden zwei Zeilen definiert: In diesem Fall „*option*“ gefolgt von den

beiden Optionen („e“ für expiration oder „i“ für inspiration). Der Benutzer definiert hier, welche Bereiche aus den entsprechend annotierten Serien extrahiert werden sollen: Hier erhält der Nutzer die Option entweder alle Intervalle, die mit „e“ oder alle Intervalle, die mit „i“ gekennzeichnet wurden zu extrahieren. Diese Option kann im Skript auch später noch um weitere Variablen ergänzt werden.

Der Befehl „*boolean*“ gefolgt von der Variablen „*timetable*“ und initialem Wert „1“ erstellt im Eingabefenster ein Feld, in dem ein Haken gesetzt werden kann. Wird der Haken gesetzt, wird der Wert „1“ in die Variable „*timetable*“ gespeichert, falls nicht, der Wert „0“. Vor dem Codeblock, der eine Tabelle mit Zeitwerten der einzelnen Tiers erstellt (siehe 3.1.1.7), wird mit einem Conditional (*if timetable = 1*) abgefragt, ob die Erstellung der Zeitentabelle gewünscht wurde. Hat der Nutzer zu Beginn den Haken im Eingabefenster gesetzt, wird die Tabelle erstellt. Der Abschluss des Codeabschnittes, der zur Erstellung des Eingabefensters führt, wird durch „*endform*“ angezeigt (Zeile 16).

## A 1.2 Filecheck Procedure

Der Anfang einer Procedure wird durch das reservierte Wort „*procedure*“, gefolgt vom gewünschten Namen der Procedure, hier „*filecheck*“ gekennzeichnet. Es folgen beliebig viele Variablen nach dem Doppelpunkt, die nur für diese Procedure gelten. Diese Variablen sind, im Gegensatz zur üblichen Nomenklatur, zu erkennen am vorgestellten Punkt: hier „*.filename\$*“. Das Ende des Codeblocks wird durch „*endproc*“ angegeben. Soll dieser Codeblock an anderer Stelle wieder aufgerufen werden, geschieht dies durch ein „@“, gefolgt vom Namen der Procedure sowie den Variablen, die je nach Codeabschnitt abweichend angepasst werden können (hier z.B. „*@filecheck: .nameofwavfile\$*“) (vgl. Kapitel 3.1.1.3).

Bevor eine Sounddatei oder eine Ergebnistabelle gespeichert wird, überprüft man mit dieser Procedure, ob eine Datei gleichen Namens bereits existiert. Anschließend erhält der Nutzer in einem Auswahlfenster die Möglichkeit, die Datei zu überschreiben oder die Ausführung des Skriptes abubrechen.

Mit einem Conditional innerhalb der Procedure wird zunächst vor einem Speichervorgang überprüft, ob bereits eine Datei mit demselben Dateinamen existiert (`if fileReadable (.filename$)`). Stellt sich diese Bedingung als wahr heraus, wird das Skript mit dem Befehl `„beginPause:“` pausiert und öffnet gleichzeitig ein Auswahlfenster (Abbildung 19). Analog zu den Befehlen in der Eingabeform (Kapitel 3.1.1.1) kann mit dem Befehl `„comment:“` Text in dieses Eingabefenster ausgegeben werden. In diesem Fall handelt es sich um Warnungen, dass die Datei bereits existiert, ihr Dateipfad- und Name sowie die Frage, ob die Datei überschrieben werden soll. Über den Befehl `„endPause:“` definiert man die verschiedenen Schaltflächen, hier `„Cancel“` sowie `„Continue & overwrite“`. Je nachdem, auf welche Schaltfläche der Nutzer klickt, wird in die Variable `„clicked“` eine Zahl gespeichert. Klickt der Nutzer auf `„Cancel“`, werden alle Objekte entfernt und das Skript terminiert. Klickt er auf `„Continue & overwrite“`, wird die Pause beendet und die Ausführung des Skriptes fortgesetzt.

## A 1.3 Äußerer Loop

Während `„for“` den Beginn der Schleife anzeigt, zeigt `„endfor“` das Ende an. `„for“` wird gefolgt von der zugehörigen Zählervariablen, hier `„ifile“`. Der Zähler beginnt bei 1 und erhöht sich pro Durchgang der Schleife um den Wert 1. Er endet schließlich bei einer ebenfalls in der ersten Zeile der Schleife festgelegten Maximalanzahl an Durchgängen. Hier wird wieder auf die zuvor gespeicherte Anzahl an Strings in `„numberOfFiles“` zurückgegriffen.

Die Codezeilen, die sich zwischen `„for“` und `„endfor“` befinden, werden so oft wiederholt, bis die Maximalanzahl erreicht ist, welche hier der Anzahl der durch das Programm aufgefundenen Sounddateien im Stringsobjekt entsprechen.

Direkt im Anschluss wird der Name der zu bearbeiteten ersten Sounddatei in eine Stringvariable geschrieben und anschließend mit dem Befehl `„Read from file:“` in das PRAAT-Objects-Window geladen. Dateipfad und -name werden aus zuvor gespeicherten Stringvariablen entnommen und so zum kompletten Dateipfad zusammengefügt.



Der Dateiname der nun im PRAAT Objects Window geöffneten und automatisch markierten Datei wird zusätzlich zur späteren Verwendung in einer Stringvariablen („*soundname\$*“) gespeichert.

## A 1.4 Korrespondierendes Textgridfile laden

Zunächst werden der Dateipfad und der Dateiname für das Textgridfile aus anderen Variablen zusammengesetzt: aus zuvor im Eingabefenster definiertem Dateipfad (*textGrid\_directory\$*), Namen der Sounddatei (*soundname\$*) und im Eingabefenster definierter Dateiendung (*textGrid\_file\_extension\$*). Der zusammengesetzte Term (Dateipfad und -name) wird anschließend in die Stringvariable „*gridfile\$*“ gespeichert. Da die Dateiendung für Textgrids im Eingabefenster (Abbildung 18) frei wählbar ist, erlaubt das Skript eine einfache Anpassung durch den Nutzer, falls die Dateiendungen der Textgridfiles vom üblichen „TextGrid“ abweichen sollten. Da die Sounddateien ausschließlich als .wav-Dateien vorliegen, wurde auf ein entsprechendes Feld in der Eingabeform (Abbildung 18) verzichtet und die Suche nach Sounddateien im Skript auf .wav-Dateien beschränkt.

Mit einem Conditional, gekennzeichnet durch das vorangestellte reservierte Wort „*if*“ und dem Befehl „*fileReadable (gridfile\$)*“, wird die Lesbarkeit der Textgriddatei geprüft.

Existiert eine passende Datei, die den Kriterien entspricht (korrekter Dateipfad, richtiger Dateiname, richtige Endung), wird Sie mit dem Befehl „*Read from file:*“ ebenfalls in das PRAAT Objects Window geladen. Das if-Conditional verbessert die Stabilität des Skriptes; findet sich keine Datei (falscher Dateipfad, beschädigte Datei) wird die Verarbeitung der Sounddatei übersprungen und endet so nicht in einer Fehlermeldung oder einem Absturz des Skriptes.

## A 1.5 Zeitentabelle

Ob der Haken für die Erstellung der Tabelle gesetzt wurde, wird zunächst mit einem Conditional überprüft (wie in Kapitel 3.1.1.1 beschrieben): wurde der Haken in der Eingabeform gesetzt, enthält die Variable „*timetable*“ den Wert 1, andernfalls den Wert 0, folglich wird mit „*if timetable =1*“ die „Wahrheit“ der Bedingung geprüft. Das Ende des Codeblocks wird mit „*endif*“ angezeigt.

Zunächst wird nun das aktuelle TextGrid per „*selectObject*“ angewählt und damit aktiv geschaltet. Anschließend wird die Anzahl der Intervalle in Tier 1 abgerufen „*Get number of intervals: 1*“ und simultan in die Variable „*numberOfIntervals*“ gespeichert. Ein Tabellenobjekt wird mit dem Befehl „*Create table with column names:*“ erstellt. Auch der Tabelle wird eine Variable („*table\_time*“) zugewiesen.

Mithilfe eines Loops wird nun iterativ für jedes Tier Startzeitpunkt, Endzeitpunkt und Name abgerufen (*Get start point/end point/ label of interval*) und in eine dafür vorgesehene (String-) Variable gespeichert (*startTime, endTime, name\$*). Die Intervalldauer wird aus Start- und Endzeitpunkt berechnet und ebenfalls in eine Variable gespeichert „*duration = endTime - startTime*“. Mit „*Append row*“ wird eine Zeile in der Tabelle angefügt und die zuvor gespeicherten Werte in die zuvor erstellte Tabelle gespeichert: „*Set numeric/string value: current\_row, ...*“. Pro Intervall wird der Tabelle eine Zeile angefügt. Dies wird so lange wiederholt, bis die in „*numberOfIntervals*“ gespeicherte Anzahl an Intervallen erreicht ist: der Loop terminiert mit „*endfor*“. Um die Tabelle zur Übergabe an laufende Systeme am ZVES zu speichern, wird zunächst der Name für die Tabelle in die Variable „*timetable\_name\$*“ zusammengesetzt, mit der Filecheck Procedure (vgl. Kapitel 3.1.1.3) überprüft, ob die Datei schon vorhanden ist und schließlich als (Werte durch Tabstopp getrennt) Textdatei gespeichert. Das nicht mehr benötigte Tabellenobjekt „*table\_time*“ wird schließlich ausgewählt und entfernt.

## A 1.6 Werte in Tabelle schreiben

Der Zähler *iframe* (Zeile 156) erhöht sich in jedem Durchgang dieser Schleife um den Wert 1, bis die zuvor in „*numberOfFramesPitch*“ gespeicherte Anzahl an Frames des Pitchobjektes erreicht ist. Wurde ein Zeitwert in die Variable „*time*“ und ihr zugehöriger Pitchwert in die Variable „*pitchVal*“ gespeichert, kann der Wert anschließend in die Tabelle gespeichert werden: per „*selectObject:*“ wird wieder das Tabellenobjekt ausgewählt, per „*Append row*“ wird eine neue Zeile in die Tabelle eingefügt. Die aktuelle Anzahl der Zeilen wird per Befehl „*Get number of rows*“ in die Variable „*current\_row*“ gespeichert. Mit dem Befehl „*Set numeric value:*“ und den zugehörigen Argumenten (getrennt durch Kommata): Reihe in die der Wert geschrieben werden soll (hier aus der eben angelegten Variable „*current\_row*“), Name der Spalte in Anführungszeichen (*time* beziehungsweise *pitch*), und dem Wert, der in die Tabelle geschrieben werden soll (aus den Variablen „*time*“ und „*pitchVal*“), wird dann pro Zeile ein Zeit- sowie ein Pitchwert in die Tabelle geschrieben bis die in „*numberOfFramesPitch*“ (vgl. Codeabschnitt 13) hinterlegte Maximalanzahl an Frames erreicht ist.

## A 1.7 Werte in Textdateien abspeichern

Für die fo-Werte werden zunächst die jeweiligen Tabellenobjekte durch den Befehl „*selectObject*“ erneut ausgewählt. In der Variablen „*nameofpitchfile\$*“ wird der gewünschte Speicherort aus Dateipfad (*new\_output\_directory\$*), Dateiname (*nameofsound\$*) und der im Eingabefenster angegebenen Endung (*output\_file\_extension\$*) zusammengesetzt. Mit der *filecheck*-Prozedure (*@filecheck: nameofpitchfile\$* vgl. Kapitel 3.1.1.3) wird sichergestellt, dass keine Datei versehentlich überschrieben wird. Anschließend wird die Tabelle, mit den Spalten durch Tabstopps separiert, als Textdatei gespeichert (*Save as tab-separated file: nameofpitchfile\$*). An die Datei wird per Befehl „*appendFileLine:*“ eine Fußzeile angehängt, die den Dateinamen und die zur fo-Analyse verwendeten Einstellungen enthält. Konnte kein Wert für „*ceiling*“

(z. B. wegen noise) ermittelt werden (vgl. Kapitel 0), erhält die Fußzeile den Hinweis *"Standard settings because Ceiling undefined"*. Das Vorgehen für die Intensitätsdateien gestaltet sich hierzu analog. Zusätzlich wird das extrahierte Soundobjekt noch per Befehl *„Save as WAV file:“* unter Angabe von gewünschtem Verzeichnis, Soundname und Dateiendung gespeichert.

Die nur vorübergehend für die Analyse benötigten Pitch-, Intensitäts- und Tabellenobjekte sowie das soeben als WAV-File gespeicherte und zuvor aus der Langserie extrahierte Soundobjekt werden nacheinander per *„selectObject“* ausgewählt und sukzessive aus dem PRAAT Objects Window entfernt (Codeabschnitt 18). Der „innere“ Loop für die Pitchanalyse wird schließlich mit *„endfor“* abgeschlossen und die Analyseroutine für das nächste Soundfile kann beginnen.

## A 1.8 Überprüfen und/oder ändern der Voreinstellungen durch den Benutzer

Zunächst werden mit dem Befehl *„selectObject“* die aktuelle Sounddatei und das aktuelle Textgrid ausgewählt. *„View&Edit“* öffnet anschließend das Editorfenster, das benötigt wird, um die Kontrolle der Annotation vorzunehmen.

*„Comment“* gefolgt von einem Text resultiert in einer Textzeile im Eingabefenster, *„real“* erstellt zwei Eingabefelder für reelle Zahlen im Auswahlfenster, die beiden Argumente in Anführungszeichen dienen der Feldbeschriftung respektive dem Standardwert, der zunächst in das Feld geschrieben wird.

*„endPause:“* gefolgt von den Namen der einzelnen Schaltflächen in Anführungszeichen gibt abhängig von der durch den Benutzer geklickten Schaltfläche einen numerischen Wert zurück, der sogleich in die Variable *„clicked“* gespeichert wird. Dies ermöglicht nun, je nach geklickter

Schaltfläche und unter Zuhilfenahme eines Conditionals, die Ausführung unterschiedlicher Befehle.

Klickt der Benutzer auf die Schaltfläche *Cancel*, wird der Codeblock nach `„if clicked=1“` ausgeführt: aktueller Sound und Textgrid werden entfernt `„removeObject“` und das Skript wird per `„exitScript“` beendet.

Klickt der Benutzer jedoch auf `„Try different settings“` wird der Codeblock nach `„elsif clicked =3“` ausgeführt: das aktuelle TextGrid wird ausgewählt (`„selectObject“`) und entfernt (`„Remove“`). Ein neues Textgrid mit den zuvor im Eingabefenster angegebenen geänderten Werten, die in `„pitch“` und `„threshold“` gespeichert wurden, wird erstellt.

Der komplette Codeblock zur Überprüfung wird eingerahmt durch einen repeat-Loop: Der Codeblock wird so lange wiederholt bis der Benutzer auf `„Continue“` klickt, also bis `„endPause“` den Wert 2 retourniert.: `„until clicked=2“` (vgl Code Annotation 7) .

## A 1.9 Nummerierung der Intervalle

Per `„selectObject“` wird das Textgrid zunächst in der Objektliste ausgewählt. Die Anzahl der automatisch erstellten und gegebenenfalls vom Nutzer weiter präzisierten Intervalle der Langserie wird in die Variable `„numberOfIntervals“` gespeichert. Mit `„schrei_no = 0“` und `„leise_no = 0“` werden zwei Variablen initialisiert. Mithilfe eines Loops werden nun die Intervalle nacheinander in aufsteigender Reihenfolge durchnummeriert. Das jeweils aktuelle Intervall wird in die Stringvariable `„label$“` gespeichert. Ob es sich nun um ein „stummes“ „i“ oder Laute enthaltendes Intervall „e“ handelt, wird mit einem Conditional `„if label$ =“` geprüft. Die zuvor initialisierten Variablen `„schrei_no“` und `„leise_no“` werden pro Treffer um den Wert 1 erhöht. Der jeweilige Wert wird schließlich per Befehl `„Set interval text“` um den aktuellen Wert der Variablen `„schrei_no“` bzw. `„leise_no“` ergänzt und die Intervalle so aufsteigend durchnummeriert.

## A 2 Messwerte der ermittelten Differenzen der Vokalisationslängen (Lges [s]) zwischen SAM und MM-Methode auf Signallevel

In den folgenden Tabellen (Tabelle 48 bis Tabelle 54) sind die Differenzen der Vokalisationslängen (Lges [s]) auf Signallevel zwischen den verschiedenen Vermessungsmethoden (SAM- bzw. MM-Methode) berichtet.

Tabelle 48: Differenzen der gemessenen Vokalisationslängen Lges [s] auf Einzelsignallevel zwischen SAM und MM-Methode der Strukturkategorie 1B

Gruppe	Pb	Signal	Diff_L1B
LKGS	1	AA08205m	0,078
LKGS	1	AA08206m	0,078
LKGS	1	AA08224c	0,068
LKGS	1	AA08226c	0,043
LKGS	3	AC05208c	0,048
LKGS	3	AC05209c	0,060
LKGS	3	AC05210c	0,059
LKGS	3	AC05213c	0,101
LKGS	3	AC05214c	0,061
LKGS	3	AC05224c	0,034
LKGS	4	AD06213c	0,013
KG	5	AA05631c	0,175
KG	5	AA05638c	0,113
KG	5	AA05639c	0,029
KG	5	AA05641c	0,057
KG	5	AA05679m	0,035
KG	5	AA05686c	0,045
KG	5	AA05678n	0,015
KG	6	AB05809m	0,006
KG	6	AB05817m	0,025
KG	6	AB05818m	0,068
KG	6	AB05821m	-0,178
KG	6	AB05823m	0,131
KG	6	AB05825c	0,127
KG	6	AB05828m	0,109
KG	6	AB05836c	0,029
KG	6	AB05840m	0,011
KG	6	AB05841m	0,022
KG	6	AB05845c	0,103
KG	6	AB05847c	0,094
KG	6	AB05850m	0,022
KG	6	AB05853m	0,020
KG	6	AB05854m	0,013
KG	6	AB05855m	0,017
KG	6	AB05856m	0,017
KG	6	AB05867n	0,080
KG	6	AB05880m	0,050

Gruppe	Pb	Signal	Diff_L1B
KG	6	AB05886m	0,100
KG	6	AB058a0c	0,151
KG	6	AB058a4m	0,040
KG	6	AB058a5m	0,037
KG	6	AB058a8m	-0,024
KG	6	AB058b1m	0,056
KG	6	AB058b3c	0,066
KG	6	AB058b4c	0,006
KG	6	AB05830c	0,140
KG	6	AB05842m	0,012
KG	6	AB05844m	0,013
KG	7	AC05601m	0,151
KG	7	AC05602m	0,034
KG	7	AC05603m	0,042
KG	7	AC05604m	0,028
KG	7	AC05610c	0,008
KG	7	AC05611c	0,020
KG	7	AC05612c	0,010
KG	7	AC05613c	0,018
KG	7	AC05614c	0,035
KG	7	AC05615c	0,012
KG	7	AC05616m	0,010
KG	7	AC05617m	0,023
KG	7	AC05618m	0,009
KG	7	AC05619m	0,013
KG	7	AC05620m	0,009
KG	7	AC05621m	0,011
KG	7	AC05622m	0,014
KG	7	AC05623m	0,006
KG	7	AC05625m	0,006
KG	7	AC05626m	-0,004
KG	7	AC05627m	0,011
KG	7	AC05629m	0,007
KG	7	AC05630m	0,036
KG	7	AC05633m	0,041
KG	7	AC05634m	0,015
KG	8	AD06902m	0,077
KG	8	AD06905m	0,052
KG	8	AD06907m	0,066
KG	8	AD06908m	0,002
KG	8	AD06913c	0,026
KG	8	AD06916c	0,120
KG	8	AD06917c	0,018
KG	8	AD06920c	0,022
KG	8	AD06927m	0,010
KG	8	AD06928m	0,046
KG	8	AD06929m	0,017
KG	8	AD06930m	0,032
KG	8	AD06937m	0,042
KG	8	AD06938m	0,052
KG	8	AD06943m	0,046
KG	8	AD06946m	0,015
KG	8	AD06953m	0,004

Gruppe	Pb	Signal	Diff_L1B
KG	8	AD06954m	0,017
KG	8	AD06963m	0,091
KG	8	AD06964m	0,002
KG	8	AD06965m	0,007
KG	8	AD06967m	0,020
KG	8	AD06971m	0,067
KG	8	AD06974m	-0,011
KG	8	AD06975m	0,020
KG	8	AD06982m	0,093
KG	8	AD06986m	0,048
KG	8	AD06994m	0,001
KG	8	AD069a3m	0,040
KG	8	AD069a4m	0,057
KG	8	AD069a6c	0,054
KG	8	AD069c0m	0,026
KG	8	AD069d3m	0,029

Tabelle 49: Differenzen der gemessenen Vokalisationslängen Lges [s] auf Einzelsignallevel zwischen SAM und MM-Methode der Strukturkategorie 1S

Gruppe	Pb	Signal	Diff_LG	Diff_B1	Diff_P	Diff_B2
LKGS	1	AA08213c	0,027	0,057	-0,039	0,009
LKGS	1	AA08221m	0,143	0,141	-0,036	0,039
LKGS	1	AA08227c	0,039	0,120	-0,131	0,051
LKGS	1	AA08217n	0,018	0,088	-0,167	0,097
LKGS	1	AA08220m	0,348	0,075		
LKGS	2	AB05302n	0,134	0,105	-0,111	0,140
LKGS	2	AB05306m	0,818	0,207		
LKGS	2	AB05314n	2,244	0,083		
LKGS	3	AC05202c	0,075	0,071	-0,021	0,025
LKGS	3	AC05216c	0,036	0,028	-0,044	0,052
LKGS	3	AC05217c	0,195	0,025	-0,033	0,203
LKGS	3	AC05218c	0,050	0,074	-0,095	0,071
LKGS	3	AC05219c	0,149	0,029	-0,031	0,151
LKGS	3	AC05227c	0,066	-1,155		
LKGS	3	AC05228c	0,470	0,461	-0,001	0,010
LKGS	3	AC05235c	-0,052	0,291	-0,052	-0,291
LKGS	4	AD06210c	0,048	0,037	-0,015	0,026
KG	5	AA05671c	0,012	0,082	-0,105	0,035
KG	6	AB05803m	0,013	0,051	-0,069	0,031
KG	6	AB05812c	0,013	0,155	-0,209	0,067
KG	6	AB05834m	-0,004	-0,009	-0,023	0,028
KG	6	AB05846c	0,096	0,005		
KG	6	AB05848c	0,109	0,011		
KG	6	AB05860n	0,007	0,005	-0,011	0,013
KG	6	AB05862n	0,098	0,118	-0,087	0,067
KG	6	AB05873c	0,025	-0,017	0,088	-0,046
KG	6	AB05875c	0,038	-0,103	0,149	-0,008
KG	6	AB05889m	0,177	0,073		
KG	6	AB05890c	0,017	0,056	-0,043	0,004
KG	6	AB058b2c	0,029	0,077	-0,074	0,026



Gruppe	Pb	Signal	Diff_LG	Diff_B1	Diff_P	Diff_B2
KG	6	AB058b6c	0,021	0,040	-0,028	0,009
KG	6	AB05826m	0,044	0,011	0,043	-0,009
KG	6	AB058b5c	0,004	-0,005	0,028	-0,019
KG	6	AB05839c	0,019	-0,039	0,059	-0,001
KG	8	AD06910m	0,026	0,061	-0,052	0,018
KG	8	AD06968c	0,008	0,014	0,034	-0,040
KG	8	AD06973m	0,564	0,065		
KG	8	AD06995m	0,058	0,047	-0,009	0,019
KG	8	AD069a0c	0,095	0,186	-0,106	0,015
KG	8	AD069c9m	0,337	-0,139	-0,094	0,570

Tabelle 50: Differenzen der gemessenen Vokalisationslängen Lges [s] auf Einzelsignallevel zwischen SAM und MM-Methode der Strukturkategorie 2B

Gruppe	Pb	Signal	Diff_L2B
LKGS	3	AC05201c	0,030
LKGS	3	AC05211c	0,033
LKGS	3	AC05221c	0,025
LKGS	3	AC05222c	0,020
LKGS	4	AD06206c	0,039
LKGS	4	AD06207c	0,030
KG	5	AA05642c	0,057
KG	5	AA05647c	0,044
KG	5	AA05650c	0,024
KG	5	AA05651m	0,022
KG	5	AA05652c	0,045
KG	5	AA05656c	0,057
KG	5	AA05659c	0,085
KG	5	AA05660c	0,058
KG	5	AA05670c	0,024
KG	5	AA05675c	0,042
KG	5	AA05681c	0,033
KG	5	AA05682c	0,088
KG	5	AA05657u	0,103
KG	6	AB05815m	0,066
KG	6	AB05837c	0,037
KG	6	AB05851m	0,015
KG	6	AB05852m	0,010
KG	6	AB05870m	0,058
KG	6	AB058b7c	0,054
KG	7	AC05632m	0,009
KG	7	AC05637m	0,018
KG	8	AD06903m	0,047
KG	8	AD06904m	0,036
KG	8	AD06914c	0,029
KG	8	AD06932m	0,142
KG	8	AD06949m	0,015
KG	8	AD06959m	0,067
KG	8	AD06984c	0,123
KG	8	AD06992m	0,183
KG	8	AD06998m	0,013

Gruppe	Pb	Signal	Diff_L2B
KG	8	AD06999c	0,012
KG	8	AD069b7c	0,444
KG	8	AD069c3m	0,184
KG	8	AD069d4c	0,023

Tabelle 51: Differenzen der gemessenen Vokalisationslängen Lges [s] auf Einzelsignallevel zwischen SAM und MM-Methode der Strukturkategorie 2S

Gruppe	Pb	Signal	Diff_LG	Diff B1	Diff P12	Diff B2	Diff P23	Diff B3
LKGS	1	AA08207m	0,151	0,147	-0,043	0,050	-0,077	0,074
LKGS	1	AA08208c	0,033	0,127	-0,151	0,058	-0,042	0,042
LKGS	1	AA08209n	0,682	-0,284	-0,454	0,409		
LKGS	1	AA08212c	1,484	0,059	-0,055	0,103		
LKGS	1	AA08214c	1,604	-0,671				
LKGS	1	AA08222m	1,599	0,316				
LKGS	2	AB05308m	0,883	0,187	-0,190	0,364		
LKGS	2	AB05309m	0,833	0,068	-0,125	0,301		
LKGS	2	AB05318c	0,588	0,149	-0,007	0,110		
LKGS	3	AC05223c	0,066	0,018	-0,015	0,056	-0,116	0,123
LKGS	3	AC05226c	0,044	0,004	0,027	0,178	-0,515	0,350
LKGS	3	AC05230c	0,040	0,026	-0,026	0,044	-0,050	0,046
LKGS	3	AC05231c	0,034	0,024	-0,018	0,044	-0,073	0,057
KG	6	AB05819c	0,672	0,048	-0,009	0,059		
KG	6	AB05864c	0,062	0,054	-0,006	0,029	-0,038	0,022

Tabelle 52: Differenzen der gemessenen Vokalisationslängen Lges [s] auf Einzelsignallevel zwischen SAM und MM-Methode der Strukturkategorie 3B/MB

Gruppe	Pb	Signal	QSS	Diff_L
LKGS	3	AC05215c	3B	0,032
LKGS	4	AD06202c	3B	0,078
LKGS	4	AD06204c	3B	0,063
LKGS	4	AD06212c	3B	0,015
KG	5	AA05629c	3B	0,020
KG	5	AA05653c	3B	0,043
KG	5	AA05674c	3B	0,028
KG	5	AA05677c	3B	0,037
KG	6	AB05888m	3B	0,012
KG	8	AD06918c	3B	0,011
KG	8	AD06922c	3B	0,091
KG	8	AD069b5c	3B	0,044
KG	8	AD069c7m	3B	0,064
KG	8	AD069d0c	3B	0,034
LKGS	3	AC05207c	MB	0,103
KG	6	AB05891c	MB	0,017
KG	6	AB058a1c	MB	-0,069

Tabelle 53: Differenzen der gemessenen Vokalisationslängen Lges [s] auf Einzelsignallevel zwischen SAM und MM-Methode der Strukturkategorie 3S

Gruppe	Pb	Signal	Diff LG	Diff B1	Diff P12	Diff B2	Diff P23	Diff B3	Diff P34	Diff B4
LKGS	1	AA08218c	0,426	0,057	-0,089	0,139	-0,112	0,119		
LKGS	2	AB05307n	2,072	0,146	-0,153	0,202				
LKGS	2	AB05319c	1,513	-0,670	-0,523	0,383	-0,138	1,394		
LKGS	3	AC05220c	0,037	0,027	-0,022	0,020	-0,045	0,049	-0,053	0,061
LKGS	3	AC05229c	0,060	-1,333						
KG	6	AB05814u	-0,049	0,079	-0,070	0,024	0,102	-0,130	0,011	-0,066
KG	6	AB05877m	0,135							
KG	8	AD069d2c	1,100							

Tabelle 54: Differenzen der gemessenen Vokalisationslängen Lges [s] auf Einzelsignallevel zwischen SAM und MM-Methode der Strukturkategorie KS

Gruppe	Kind	Signal	Diff LG	Diff B1	Diff P	Diff B2
LKGS	1	AA08225c	0.022	0.015	-0.019	0.027
LKGS	2	AB05303m	1.246	0.838		
LKGS	3	AC05212c	0.035	0.045	-0.083	0.073
LKGS	3	AC05225c	0.110	0.021	-0.057	0.146
LKGS	3	AC05233c	0.077	0.016	-0.017	0.078
LKGS	4	AD06211c	0.020	0.012	-0.010	0.018
KG	5	AA05626c	0.051	0.032	-0.041	0.060
KG	5	AA05663c	0.031	-0.039	0.083	-0.013
KG	5	AA05673c	-0.544	-0.014	0.071	-0.601
KG	6	AB05887m	0.030	0.103	-0.096	0.023
KG	6	AB05827m	0.011	0.050	-0.072	0.033
KG	6	AB05831c	-0.631	0.088	-0.059	-0.659
KG	7	AC05605m	0.041	0.022	-0.064	0.083
KG	7	AC05636m	0.037	0.048	-0.062	0.051
KG	8	AD069b2m	0.317	0.297	-0.063	0.083

