

Supplementary Information

All supplementary files belong to:

Voice Features in Paediatric ADHD: Baseline Group Differences and Longitudinal Change Under Methylphenidate. European Child & Adolescent Psychiatry. Bamberger, R., Kuhles, G., Lotter, L.D., Dukart, J., Konrad, K., Günther, T., Siniatchkin, M., Fuchs, M., & Von Polier, G.
Corresponding author: Rachel Bamberger (Research Centre Jülich, INM 7), r.bamberger@fz-juelich.de

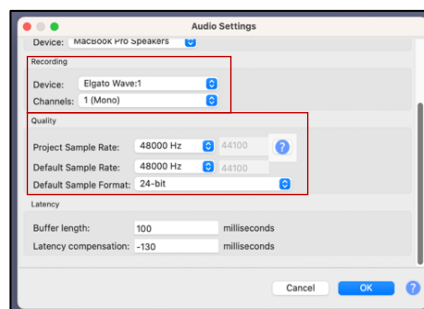
Appendix S1

SOP: Spontanspracherhebung

Material	Aufgaben
- Instruktionsanweisung - Bild <i>Cookie Theft Task</i> (neue Version) - Stoppuhr/Handy - Headset-Mikrofon; Verstärker; Adapter - Laptop (mit USB-Anschluss) - Lineal (Abstandshalter; 2cm)	- Namen sagen, Zählen (1-10), - Bildbeschreibung/Nacherzählung/gelenkte Spontansprache: <i>Cookie Theft Task</i> - Freie Spontansprache: Vergangenes Ereignis und Traumtag beschreiben - Vokalhalteversuche: /a/ halten (normal/laut)

Setup und Aufnahme

- Das Headset wird in einem Abstand von 2cm zum Mundwinkel des Kindes positioniert
- Headset über Adapter an Verstärker (Scarlett Solo) und dann an Laptop anschließen
- Programm *Focusrite Control 2* öffnen (nur beim ersten Mal oder falls Probleme)
 - o In den Einstellungen *Preferences* → *Combine Input 1 and Input 2* auswählen
- Programm *Audacity* öffnen
- Aufnahmeeinstellungen bei Audacity
 - o *Audio Setup* → *Recording Device* → *Scarlett Solo 4th Gen* auswählen (wenn Headset nicht gefunden wird: *Audio Einrichtung* → *Audiogerät erneut suchen* → sollte jetzt als Option bei *Aufnahmegerät* zu finden sein)
 - o Unter *Audioeinstellungen* Parameter korrekt definieren
 - Aufnahmekanäle: 1 (Mono)
 - Qualität:
 - Sample Rate (Abtastrate): 48000 Hz
 - Sample Format (Abtastformat): 24-bit



- Die Stimmufnahmen sollen bei geschlossenem Fenster gemacht werden
- Zur Dokumentation der Umgebungsbedingungen soll bei jedem Kind eine Aufnahme von 5 Sekunden ohne Phonation gemacht werden (gesondert abspeichern)
- Die einzelnen Aufgaben sollen jeweils als eigene Aufnahmen gemacht werden

Audacity: Mehrere Aufnahmen in derselben Spur erstellen

- Erste Aufnahme durchführen und darauf achten, dass man die Aufnahme stoppt (roter Knopf zum Starten und Stopp-Knopf (quadratisch) zum Beenden)
- Neue Aufnahme wird gestartet, in dem man wieder auf den Aufnahmeknopf klick (roter Knopf)
- Falls während der Aufnahme Prompts gegeben werden müssen: Aufnahme pausieren und wieder starten
- Prozedur wiederholen für weitere Aufnahmen
- Speichern der Spuren: gewünschte Spur auswählen über *Strg + Klick*
- *Datei* → *Exportieren* → *Ausgewähltes Audio exportieren...* → Als WAV-Datei speichern
- Korrekt benennen (siehe Nachbereitung)

Instruktionen

„Ich werde dir jetzt ein paar Fragen stellen oder Anweisungen geben, auf die du antworten sollst bzw. die du befolgen sollst. Weil wir davon eine Aufnahme machen, sollst du jetzt das Headset aufsetzen.“

Wir fangen zuerst mit einer Aufgabe an, bei der du deinen Namen sagen und von 1-10 zählen sollst, danach habe ich ein Bild für dich, dass du beschreiben sollst, und dann stelle ich dir noch zwei Fragen, bei denen du dir etwas ausdenken musst. Am Ende sollst du ein /a/ sprechen und halten. Während der einzelnen Aufgaben mache ich dann Aufnahmen von deiner Stimme. Deshalb spreche ich währenddessen nicht so viel. Hast du noch Fragen? (...) Gut, dann fangen wir an.

Ich möchte nur deine Stimme hören, deshalb werde ich während der Aufnahmen nicht mit dir sprechen. Versuch also so lange zu erzählen, bis ich Stopp sage. Wenn ich dieses Zeichen mache, bedeutet es, dass du noch ein wenig erzählen musst.“

Mittlere Sprechstimmlage (ungespannt)/Indifferenzlage
„Bitte sag mir wie du heißt, nur deinen Vornamen, und zähle dann ganz entspannt von 1-10.“
Bildbeschreibung
„Du hast jetzt ein wenig Zeit, genauer gesagt 90 Sekunden, um so ausführlich wie möglich zu erzählen, was auf dem Bild passiert bzw. mir das Bild zu beschreiben. Erzähl mir das so genau wie möglich. Versuch die ganze Zeit über zu sprechen. Wenn dir nichts mehr einfällt, kannst du beschreiben, was du siehst. Du kannst keine Fehler machen und darfst mir alles erzählen, was du möchtest.“
Freie Spontansprache
„Ich stelle dir jetzt zwei Fragen, auf die du mir etwas ausführlicher antworten sollst. Du hast für jede Frage eine Minute (60 Sek.) Zeit. Bei dieser Aufgabe gibt es auch kein richtig oder falsch und du kannst alles erzählen, was du möchtest.“ „Beschreib‘ mir bitte etwas Schönes, das du schon erlebt hast, wie z.B. deinen letzten Urlaub oder deine letzte Klassenfahrt.“ „Und jetzt überleg mal, wie dein schönster Tag aussehen würde, wenn du machen könntest, was du willst und Geld oder ähnliches keine Rolle spielen würden.“ (Auch bei kurzen Pausen Zeit geben, sodass der Proband:in wieder etwas einfallen kann. Sollte ihm/ihr ca. 15 Sekunden lang trotzdem nichts einfallen, sollten <u>offene</u> Fragen zu bereits Erzähltem gestellt werden bzw. einzelne Aspekte aufgegriffen werden, dafür soll die Aufnahme pausiert werden, damit der/die Untersucher:in nicht zu hören ist)
Vokalhalteversuche
„Du hast jetzt gleich die Aufgabe, den Vokal /a/ für etwa 5 Sekunden in einer für dich angenehmen Lautstärke und Tonhöhe zu halten. Ich zeige dir einmal, wie das geht (Vormachen der drei Versuche nacheinander). Bevor wir mit der Audioaufnahme starten, darfst du jetzt noch einmal üben. Du kannst die drei Versuche nacheinander machen, mit einer kurzen Pause dazwischen. Danach beginnen wir dann mit der Aufnahme. Danach sollst du das nochmal mit einer lauten Stimme machen“ „Prima! Hast du noch Fragen? (...) Okay, dann beginnen wir jetzt mit der Audioaufnahme. Bitte versuche nun, den Vokal /a/ in drei Versuchen zu halten, wie du es gerade geübt hast. Damit nur du auf der Aufnahme zu hören bist, werde ich nichts sagen. Ich gebe dir aber ein Zeichen mit der Hand, wann du aufhören kannst und wieder weitermachen sollst. (...) Prima und jetzt mach das Ganze nochmal mit lauter Stimme“ (Sollten die Übungsversuche falsch klingen (z.B. sehr hoch, laut etc.) können die Übungsversuche öfter als dreimal wiederholt werden, bis die Lautstärke und Tonhöhe für die untersuchende Person subjektiv als korrekt wahrgenommen werden.)

Appendix S2

Within each participant, the same recording equipment was used for both visits.

Initial recording setup: Use of a cardioid microphone (*Elgato Wave:1*, 24-bit/48-kHz resolution, -25–15-dBFS sensitivity) connected to a laptop with a microphone-mouth distance between 30 and 40 cm. Recordings were made using either *Voice Memos*® (Apple, 2023) (n = 6) or *Audacity*® *Version 3.0* (Audacity, 2021) (n = 4).

Final recording setup: Use of an omnidirectional headset microphone (Sennheiser HSP 2 EW 3; frequency response: 20–20,000 Hz; sensitivity: 2 mV) positioned 2 cm from the corner of the mouth to ensure a constant recording distance for each child and measurement. The headset microphone was used with a preamplifier (*Focusrite Scarlett Solo 3rd Gen*; 20 Hz – 20 kHz ± 0.1 dB frequency response, 24-bit/192-kHz resolution) and recordings were made using *Audacity*® *Version 3.0* (Audacity, 2021).

Appendix S3

For the frequency task the beginning of the first word and the end of the last word were detected, and the recording was cut according to these boundaries. For the picture description task, the retell and the fiction task the first word was detected, and the recording was cut to the desired length of 90 seconds (picture description) or 60 seconds (retell/fiction). For the phonation normal and loud tasks, the three sustained /a/ sound were detected, the middle 60% was cut out of each one (approximately 3 seconds) and these were concatenated to create a clip lasting approximately 9 second. Ambient noise recordings were checked for voice, after which they were cut 5 seconds in from the beginning.

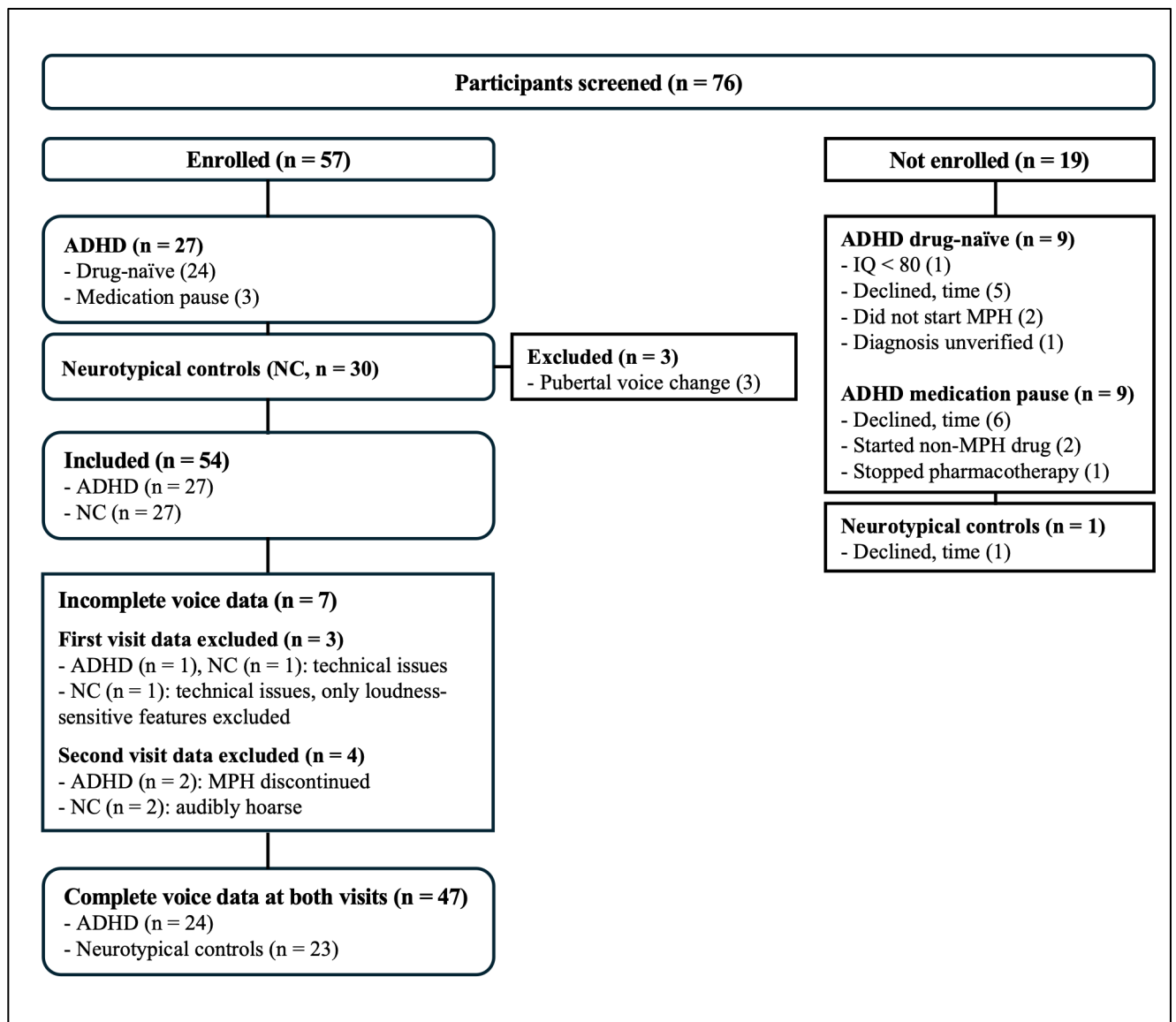
Appendix S4

The ambient noise level for measurements at a 30-cm distance should optimally be < 38 dB (non-weighted) (Patel et al., 2018). For the children who used a microphone the values in the present study ranged between 3.29 - 30.32 dB (non-weighted). With the headset microphone the maximum ambient noise level should be about 48–53 dB (non-weighted), which is considered acceptable for accurate soft voice measurements (Šrámková et al., 2015). As the children in the present study were not instructed to speak softly, we found all values acquired by the headset microphone to be acceptable, as they ranged between 41.92 - 53.82 dB (non-weighted). Values of Signal-to-Noise-Ratio (SNR) should ideally be > 10db for accurate measuring of the soundlevel in noise (Šrámková et al., 2015). SNR values measured in the present study, either by microphone or headset microphone, ranged between 20.46 - 38.95 dB.

Appendix S5

A total of 54 children who met inclusion criteria were found eligible for the study (see Figure S1 for more details). Seven of these children had missing data for various reasons: two children with ADHD dropped out after the first visit due to the termination of pharmacotherapy; two neurotypical controls were audibly hoarse due to colds at the second visit; and three children had technical issues at the first visit. Due to these issues, data from one case and one control had to be excluded entirely, and the loudness-sensitive features could not be analysed for one control that used a headset.

Figure S1 Participant flow chart from screening to final sample.



Note. ADHD = attention-deficit/hyperactivity disorder; NC = neurotypical controls; IQ = intelligence quotient; MPH = methylphenidate.

Table S1

Information on voice measurements (Equipment, task, average length)

Characteristic	Mean (SD) Length (s)*	ADHD Baseline (n)	ADHD Follow-Up (n)	Control Baseline (n)	Control Follow-Up (n)
Equipment					
Microphone	-	9	9	1	1
Headset	-	17	16	25	24
Task Completion					
Frequency	9.8 (2.8)	26	25	26	25
Picture description	89.5 (2.6)	26	25	26	25
Retell	59.9 (0.6)	24	23	26	25
Fiction	59.8 (0.7)	25	25	26	25
Phonation normal	8.5 (1.5)	20	19	26	25
Phonation loud	8.7 (1.4)	17	19	26	25

Note. Average length of the preprocessed audio recordings per task.

Table S2 Information on pharmacotreatment of children with ADHD

	Drug	Dose (mg)	Post-intake interval (hh:mm)
1	Equasym® Extended-release (ER)	50	06:45
2	Medikinet® ER	20	05:30
3	Equasym® ER	60	06:25
4	Equasym® ER	60	06:15
5	Equasym® ER	40	02:45
6	Equasym® ER	40	02:35
7	Medikinet® ER	10	02:35
	Kinecteen® Immediate-release (IR)	54	
8	Kinecteen® ER	36	04:05
9	Medikinet® ER	20	04:30
10	Medikinet® ER	10	02:30
11	Equasym® ER	30	02:50
12	Concerta® ER	36	02:15
13	Equasym® ER	30	03:50
14	Equasym® ER	20	04:00
15	Medikinet® ER	40	03:30
16	Kinecteen® ER	40	04:00
17	Equasym® ER	30	03:50
18	Equasym® ER	20	03:20
19	Ritalin LA® IR	40	03:50
20	Medikinet® ER	30	04:20
21	Ritalin LA® ER	30	03:30
22	Equasym® ER	30	03:30
23	Equasym® ER	30	03:30
24	Medikinet® ER	15	03:00
25	Ritalin LA® ER	30	03:20
Average post-intake interval (SD)			03:51 (01:13)

Note. One participant was given Medikinet® and Kinecteen® at the same time. Patients 1-22 were initially drug-naïve; 23 - 25 took a break from medication lasting ≥ 4 weeks before the follow-up. Post-intake interval is the time between taking the drug and the start of the speech assessment.

References

Apple. (2023). *Voice Memos*. In (Version 2023)

Audacity. (2021). *Audacity: Free Audio Editor and Recorder* In (Version 3.0)

Patel, R. R., Awan, S. N., Barkmeier-Kraemer, J., Courey, M., Deliyski, D., Eadie, T., Paul, D., Švec, J. G., & Hillman, R. (2018). Recommended Protocols for Instrumental Assessment of Voice: American Speech-Language-Hearing Association Expert Panel to Develop a Protocol for Instrumental Assessment of Vocal Function. *American Journal of Speech-Language Pathology*, 27(3), 887-905.
https://doi.org/10.1044/2018_ajslp-17-0009

Šrámková, H., Granqvist, S., Herbst, C. T., & Švec, J. G. (2015). The softest sound levels of the human voice in normal subjects. *Journal of Acoustical Society America*, 137(1), 407-418.
<https://doi.org/10.1121/1.4904538>