Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A longitudinal corpus of sustained phonation and voluntary cough acoustics from three ISS crew members

Gökhan Aydemir[1]

draydemir73@gmail.com

Atasehir Acibadem Hospital

Oğuzhan Aydemir[2]

İstanbul Nişantaşı Üniversitesi

data-descriptor

Keywords:

Posted Date: November 3rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10706031/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: Competing interest reported. Under the contract between the parties, TÜBİTAK UZAY and the Turkish Space Agency could derive financial benefit from sales of the software used in this investigation, as disclosed in the participant consent form. Conflict of interest disclosures were filed with the NASA Institutional Review Board for all investigators. The authors declare no other competing interests.

A longitudinal corpus of sustained phonation and voluntary cough acoustics from three ISS crew members

Gökhan Aydemir[1]* and Oğuzhan Aydemir[2]

[1] Department of Pediatrics, Acıbadem Ataşehir Hospital, Istanbul, Türkiye

[2] Nişantaşı University, Istanbul, Türkiye

ORCID: Gökhan Aydemir 0000-0001-9655-0530 · Oğuzhan Aydemir 0000-0002-1611-0947

*Correspondence: draydemir73@gmail.com

Participant identification. *The three participants are referred to exclusively as Crew M1, M2 and M3. No names, mission call-signs, agency crew codes, mission identifiers, calendar dates or file-system identifiers appear anywhere in this manuscript, its tables or its data records. Timing is reported only as mission day relative to launch. Detailed demographic characteristics are deliberately withheld because, in a crew of this size, age, sex and nationality in combination would be identifying.*

Abstract

Spaceflight perturbs several systems that participate in voice production. Systematic acoustic corpora acquired in orbit remain scarce, and they rarely include repeated within-subject sampling. We describe a corpus of 254 verified recordings from three International Space Station crew members, spanning a pre-flight quarantine phase and mission days 4 to 14. Each session followed a fixed battery administered twice daily: three voluntary coughs, then sustained /a/, /i/ and /u/. Recordings are released in two parallel forms, the original acquisition and a speech-enhanced version. Per-recording acoustic measurements, session metadata and provenance fields accompany them. Technical validation covers temporal provenance, file integrity, task-slot verification, and the effect of the enhancement stage on every measure. We also position the corpus against three public reference corpora using a single measurement pipeline. Two results are relevant to reuse. Cough timing agrees closely with a 2,645-participant corpus once burst-level segmentation is used. Absolute acoustic values order with acquisition condition, whereas dimensionless within-session vowel ratios do not. No diagnostic inference is made or supported.

Background & Summary

Human spaceflight produces a set of well-documented physiological adaptations, several of which act on structures that participate in voice production. Cephalad fluid redistribution begins within hours of orbital

insertion. Cabin environmental parameters differ from terrestrial baselines in humidity and carbon dioxide concentration. Sleep-wake structure is altered by a 90-minute orbital period. Voice is also the only physiological signal that can be sampled repeatedly, non-invasively and in seconds, using hardware already aboard. That makes it operationally attractive where consumable limits and crew time constrain sampling.

Despite this, orbital voice corpora are rare. Existing analyses of astronaut speech have generally used mission communications audio. Such material is recorded for operational purposes, with uncontrolled task and uncontrolled duration, and it has no pre-flight counterpart acquired through the same path. What has been missing is a corpus in which the same speaker performs the same standardised vocal task at scheduled intervals across a mission, with a pre-flight baseline acquired under the identical protocol.

This Data Descriptor presents such a corpus. Its design properties are:

- Fixed task battery. Every session comprises the identical six-recording sequence: three voluntary coughs, then sustained /a/, /i/ and /u/. Task duration is set by the acquisition application rather than by the participant, which makes sessions directly comparable to one another.
- Repeated within-subject sampling. Sessions were scheduled twice daily, morning and evening, giving up to 16 in-flight sessions per crew member.
- Pre-flight baseline under the same protocol. Quarantine sessions used the identical battery and the identical acquisition path. The phase contrast is therefore not confounded by a change of instrument.
- Dual release form. Both the original acquisition and the speech-enhanced version are retained. Corpora seldom publish both, which leaves reusers unable to know what an enhancement stage did to a measurement. Here that comparison is quantified.
- Complete provenance. Every record carries its acquisition-time source and precision, an integrity hash, and a task-slot verification status.

The corpus is released as derived acoustic measurements and metadata rather than audio. This follows the release model of the Oxford/UCI Parkinson's corpora, which distribute computed features with no audio and are among the most reused voice datasets in the field. The rationale is given under Data Availability.

Methods

Participants and design

Three crew members of a short-duration International Space Station mission participated. Sessions were acquired in two phases: a pre-flight quarantine phase and an in-flight phase covering mission days 4 to 14. The design is within-subject longitudinal, and each participant serves as their own reference.

The study was approved by the NASA Institutional Review Board (Federalwide Assurance No. 00019876; eIRB No. STUDY00000621, expedited review category 4) and was conducted in accordance with the Declaration of Helsinki, under protocols with the Turkish Space Agency, NASA and Axiom Space. Written informed consent for participation and for the sharing of de-identified derived data was obtained directly from all three participants; no consent waiver was sought or granted. Participants are identified only as M1, M2 and M3, and detailed demographic characteristics are deliberately withheld, because in a crew of this size age, sex and nationality in combination would be identifying.

Recording protocol

Each session consisted of six recordings administered in the fixed order set out in Table 1.

Table 1. Session task battery.

Order	Task	Instruction
1–3	Voluntary cough	Three separate coughs, one per recording
4	Sustained /a/	Comfortable pitch and loudness
5	Sustained /i/	Comfortable pitch and loudness
6	Sustained /u/	Comfortable pitch and loudness

Recording duration was fixed by the acquisition application at approximately three seconds per file, sampled at 16 kHz, single channel. Sessions were scheduled twice daily. Fixing the duration is a deliberate protocol choice with a consequence for reuse. It removes participant-controlled phonation length as a source of between-session variance, at the cost of not supporting maximum-phonation-time measures. Corpora that leave duration free carry the opposite trade-off. Vowel recordings in the Coswara corpus, for example, range from 2 to 23 seconds.

Acquisition and processing

Audio was captured by a mobile application aboard the station and transferred to ground. Two parallel versions of every recording are retained and measured: the original acquisition, and the same material

after a deep noise suppression stage (hereafter DNS64). Technical Validation reports what that stage changes and what it leaves invariant.

Acoustic measurement

All corpora reported here, the present corpus and every reference corpus alike, were processed with a single pipeline built on Praat via the Parselmouth interface. All audio was brought to 16 kHz before measurement, so that no comparison is confounded by sampling rate.

Sustained vowels. Fundamental frequency (F0), local jitter, local shimmer, harmonics-to-noise ratio (HNR) and smoothed cepstral peak prominence (CPPS).

Coughs. Two measurement conventions were computed for every cough recording, because the literature is divided between them and the choice changes reported values by an order of magnitude. Under the burst-level convention, the amplitude envelope (20 ms RMS) is segmented into individual cough bursts at 25% of peak amplitude, with a 150 ms minimum separation. Within each burst, rise time (10% to 90% of that burst's peak), duration, decay and spectral centroid are measured on a finer 5 ms envelope. The recording value is the median across bursts. Under the whole-recording convention, a single peak is located in the entire recording and the same quantities are computed from it.

Dimensionless vowel ratios. Because /a/, /i/ and /u/ are acquired seconds apart through the identical path, their ratios cancel gain, room and microphone-distance terms algebraically. HNR(/a/) ÷ HNR(/i/) is reported.

Reference corpora

The public corpora against which the present corpus is positioned are listed in Table 2.

Table 2. Public reference corpora, processed with the same pipeline.

Corpus	n	Acquisition	Tasks used here
Saarbrücken Voice Database	2,042	Sound-treated clinical room, 50 kHz	/a/, /i/, /u/
Coswara	2,709	Consumer device, home, web browser	/a/, /e/, coughs
COUGHVID	646 expert-labelled	Consumer device, crowdsourced	coughs
Oxford/UCI Parkinson's	31 / 42	Derived measures only, no audio released	— (release-model precedent)

Data Records

Table 3. Verified recordings by crew member, phase and task.

Crew	Phase	Cough	/a/	/i/	/u/	Total
M1	Quarantine	6	2	2	2	12
M1	In-flight	48	16	16	16	96
M2	Quarantine	12	4	4	4	24
M2	In-flight	21	7	7	5	40
M3	Quarantine	6	2	2	1	11
M3	In-flight	42	14	14	1	71
All		**135**	**45**	**45**	**29**	**254**

A session is defined throughout as a unique combination of crew, phase, mission day and time of day. It is therefore recomputable from the released fields alone. The totals are 2 quarantine and 16 in-flight sessions for M1, 3 and 6 for M2, and 3 and 14 for M3, giving 44 sessions. Recordings are balanced across time of day (134 morning, 120 evening).

The distributed records are derived measurements and metadata:

Table 4. Released files.

File	Content
recordings_measures.csv	One row per recording × version: crew code, phase, mission day, time of day, task, version (original / DNS64), F0, jitter, shimmer, HNR, CPPS, duration
cough_bursts.csv	264 rows — one per cough recording × version: burst-level rise, duration, decay, spectral centroid, burst count, within-recording dispersion; plus whole-recording values under both conventions
sessions.csv	44 rows — session inventory and per-session record counts
provenance.csv	254 rows — acquisition-time source and precision, truncated content hash, task-slot verification status

Metadata fields are crew (M1/M2/M3), phase (quarantine or in-flight), mission day, time of day (AM/PM), task, version (original or DNS64), timestamp source (UUID-embedded or filesystem), and the slot-verified flag.

Technical Validation

Temporal provenance

Acquisition times were not taken from filenames. Three independent sources were compared: timestamps embedded in version-1 UUID filenames, filesystem modification time, and creation time. Their agreement is reported in Table 5.

Table 5. Agreement of time sources.

Source	Records	Agreement with UUID-embedded time
UUID version 1, embedded	157	157 / 157 exact to 0.000 s
Filesystem modification time	97	median offset +0.02 h, SD 2.41 h
Creation time	—	not usable (reset during transfer)

For 157 of 254 records the acquisition instant is known to millisecond precision. The remaining 97 carry a timestamp-source flag and a stated ±2.4 h uncertainty. Analyses sensitive to time of day should use that flag.

File integrity and completeness

Every distinct recording on the acquisition media was matched by MD5 hash against the manifest. All 258 distinct source files are accounted for, and no orphaned recording exists. The expected session content of six recordings divides the corpus exactly, and an independent enumeration of source files returned the same session total.

Task-slot verification

A provenance check compared each file's naming convention against that of its own session. Sixteen /u/ records did not match their session's convention, and ten of these shared an MD5 hash with a record already assigned elsewhere. All sixteen were excluded from the verified subset. The corpus is therefore distributed as 254 verified records rather than the 270 originally enumerated, and the exclusion is recorded per record in provenance.csv. This is why the /u/ task is under-represented for M3 (Table 3). Analyses requiring a complete /a/-/i/-/u/ triad should use M1 and M2.

Effect of the enhancement stage

Because both versions are retained, the effect of speech enhancement on each measure can be quantified directly — a comparison that is rarely possible.

Table 6. Sustained vowels, median values, original versus DNS64.

Measure	Original	DNS64	Change
F0 (Hz)	100.09	101.01	+0.92
Jitter (%)	0.68	0.61	−0.07
Shimmer (%)	10.30	6.34	**−3.96**
HNR (dB)	11.90	15.26	**+3.36**
CPPS (dB)	12.09	11.53	−0.56

The enhancement stage acts strongly and specifically on the two measures defined against a noise floor, shimmer and HNR, and leaves F0 and CPPS essentially unchanged. Reusers comparing against literature values must therefore state which version they used for shimmer and HNR. For F0 and CPPS the two versions are interchangeable.

Table 7. The dimensionless ratio HNR(/a/) ÷ HNR(/i/) under the same stage.

Crew	Original	DNS64	Change
M1	0.759	0.747	−1.6%
M2	1.017	1.045	+2.8%
M3	1.038	1.022	−1.5%

Absolute HNR moves by 3.36 dB, approximately 29%, between versions, whereas the ratio moves by under 3%. The ratio is formed from two recordings made seconds apart through the same path and processed identically, so the enhancement term largely cancels. This supports its use as the cross-version and cross-corpus quantity.

Cross-corpus positioning of vowel measures

Table 8. Sustained /a/, male speakers, single pipeline, all audio at 16 kHz.

Corpus	n	F0	Jitter	Shimmer	HNR	CPPS
SVD healthy, F0 80–130 Hz (sound-treated room)	156	113.6	0.43	2.71	23.36	20.74
Coswara healthy, F0 80–130 Hz (consumer device)	577	120.7	0.61	5.31	17.62	12.90
Present corpus — original	43	98.2	0.89	10.18	11.02	**12.69**
Present corpus —	45	97.3	0.67	7.72	14.21	**11.93**

DNS64						

The measures order monotonically with acquisition condition rather than with any property of the participants. Every measure moves in the same direction from the sound-treated corpus to the two field corpora. CPPS is the measure most often used as a clinical screening threshold. It places the present corpus at 12.69 and 11.93 dB, essentially where 577 healthy consumer-device recordings sit (12.90 dB), and approximately 8 dB from the sound-treated corpus. This is a property of the measurement chain shared by field-acquired corpora generally.

Cross-corpus positioning of cough measures

Table 9. Cough timing under both measurement conventions.

Corpus	Convention	Rise (ms)	Duration (ms)	Decay (ms)	Centroid (Hz)	Bursts
Present corpus	burst-level	**11.0**	91.4	27.1	715	2
Coswara (n = 2,645)	burst-level	**9.9**	186	39	1051	4
Present corpus	whole-recording	108.9	276.8	33.7	765	—
Coswara	whole-recording	283	710	62	1092	—
COUGHVID (n = 646)	whole-recording	380	823	83	1149	—

Two points follow. First, the measurement convention rather than the corpus drives the reported values. The whole-recording convention treats inter-burst silence as part of the rise, and it inflates rise time by roughly an order of magnitude in every corpus containing multiple coughs per recording. Under burst-level segmentation, the present corpus and a 2,645-participant consumer corpus agree to within 1.1 ms. Cough timing values should not be compared across studies without stating the convention. Second, burst count reflects protocol. The present protocol requests one cough per three-second recording, with a median of 2 detected bursts, whereas Coswara requests a cough sequence in a longer recording, with a median of 4. Duration differences follow directly.

Spectral centroid varies substantially between the three crew members (medians 570, 1422 and 768 Hz), spanning the range of both large reference corpora. With three participants, corpus-level centroid comparison is not informative and none is drawn.

Dimensionless ratio across corpora

Table 10. HNR(/a/) ÷ HNR(/i/).

Corpus	Median	Within-corpus CV
SVD healthy (sound-treated)	0.874	14.8%
Coswara healthy (consumer device)	0.926	51.1%
Present corpus, original	0.759 / 1.017 / 1.038 (per crew)	—
Present corpus, DNS64	0.747 / 1.045 / 1.022 (per crew)	—

Corpus medians span 0.874 to 0.926 across acquisition conditions that displace absolute CPPS by approximately 8 dB. Per-crew values in the present corpus fall within and around that band. The wider dispersion in Coswara is consistent with its free task duration relative to the fixed durations of the other two corpora.

Usage Notes

Recommended comparisons

- Within-subject, within-phase and phase-contrast analyses are the designed use of this corpus and require no cross-corpus normalisation.
- Cross-corpus comparisons should use dimensionless within-session ratios (Table 10), which are stable across acquisition condition and across the enhancement stage.
- Cough analyses must state the segmentation convention; burst-level values are directly comparable to Coswara (Table 9).

Absolute thresholds

Clinical screening thresholds for CPPS, HNR and shimmer in the published literature are derived from sound-treated recordings. Table 8 shows that absolute values shift systematically with acquisition condition across all field corpora examined, including a corpus of 577 healthy participants recorded on consumer devices. Absolute thresholds from sound-treated corpora should therefore not be applied to this corpus, nor to any field-acquired corpus, without re-derivation of norms under the matching acquisition condition. This is a property of the measurement chain, and it carries no implication about the participants.

Version selection

The original and DNS64 versions may be used interchangeably for F0 and CPPS. The version must be stated explicitly for shimmer and HNR (Table 6). Ratio quantities may be taken from either (Table 7).

Design characteristics for secondary analysis

Table 11. Sessions available and smallest within-subject effect detectable at 80% power, α = 0.05, for the phase contrast.

Crew	Quarantine sessions	In-flight sessions	Detectable effect
M1	2	16	$d \geq 2.24$
M2	3	6	$d \geq 2.31$
M3	3	14	$d \geq 1.91$

Two-sample t-test, two-sided. Sessions defined as in Table 3.

The in-flight series of 6 to 16 sessions per crew member supports within-phase trajectory analysis. The phase contrast is sensitive to large effects only, because the pre-flight window afforded fewer scheduled sessions than the flight window. Investigators designing follow-on missions can use Table 11 directly. At 80% power and α = 0.05, nine sessions per phase per participant reach d = 1.5, eleven reach d = 1.3, and seventeen reach d = 1.0. Balancing the pre-flight window against the flight window is the single most effective design change available.

Mean session clock time differs between phases by 2.8 to 8.1 hours, depending on crew member. Phase contrasts and time-of-day effects are therefore not orthogonal in this corpus. Time-matched subsets should be used where circadian effects are of interest.

Not supported by this corpus

This Data Descriptor supports no diagnostic claim, no disease classification and no individual clinical interpretation. The corpus is released for methodological work on acoustic measurement under spaceflight conditions, for protocol design of follow-on missions, and as a cross-corpus reference point for field-acquired voice.

Data Availability

The derived acoustic measurement tables and metadata described in Table 4 are deposited at Zenodo under the identifier 10.5281/zenodo.21939101 (https://doi.org/10.5281/zenodo.21939101). They are released under a CC BY 4.0 licence and are publicly downloadable without restriction. These files contain no audio, and no direct or indirect participant identifiers beyond the crew codes M1, M2 and M3.

Raw and enhanced audio recordings are not distributed. Voice and cough recordings are biometric data from which speaker identity can be recovered, and the participants are three members of an identified professional cohort. Public release of the audio could not be reconciled with the confidentiality

commitments given in the study consent. This release model, in which computed acoustic features are distributed publicly and source audio is withheld, follows that of the Oxford/UCI Parkinson's voice corpora. Those corpora contain no audio and are among the most reused voice datasets in the field. Requests for access to audio for verification purposes will be considered by the corresponding author, subject to the approvals governing the study.

Code Availability

All measurement code used to produce the released tables and every cross-corpus comparison reported here is deposited with the data at Zenodo under the identifier 10.5281/zenodo.21939101 (https://doi.org/10.5281/zenodo.21939101). It comprises the sustained-vowel measurement script and the voluntary-cough measurement script. Both are self-contained command-line programs implemented in Python 3, using Praat 6.x through the Parselmouth interface with the parameters stated in Methods, and both are accompanied by a dependency specification and usage documentation. The identical code and parameters were applied to the public reference corpora. Every cross-corpus comparison in this Data Descriptor is therefore reproducible from public data plus the released tables. No custom or proprietary components are required.

References

1. Sharma, N. K. et al. Coswara: A respiratory sounds and symptoms dataset for remote screening of SARS-CoV-2 infection. Sci. Data 10, 397 (2023).

2. Orlandic, L., Teijeiro, T. & Atienza, D. The COUGHVID crowdsourcing dataset, a corpus for the study of large-scale cough analysis algorithms. Sci. Data 8, 156 (2021). Dataset: https://zenodo.org/records/7024894

3. Saarbrücken Voice Database. Institute of Phonetics, Saarland University. Dataset: https://zenodo.org/records/16874898

4. Cesari, U. et al. A new database of healthy and pathological voices. Comput. Electr. Eng. 68, 310–321 (2018). Dataset: https://physionet.org/content/voiced/

5. Little, M. A., McSharry, P. E., Hunter, E. J. & Ramig, L. O. Suitability of dysphonia measurements for telemonitoring of Parkinson's disease. IEEE Trans. Biomed. Eng. 56, 1015–1022 (2009).

6. Tsanas, A., Little, M. A., McSharry, P. E. & Ramig, L. O. Accurate telemonitoring of Parkinson's disease progression by noninvasive speech tests. IEEE Trans. Biomed. Eng. 57, 884–893 (2010).

7. Boersma, P. & Weenink, D. Praat: doing phonetics by computer, version 6.x. http://www.praat.org

8. Jadoul, Y., Thompson, B. & de Boer, B. Introducing Parselmouth: A Python interface to Praat. J. Phon. 71, 1–15 (2018).

Author Contributions

G.A. conceived the study, designed the investigation, obtained the regulatory and ethical approvals, and led data collection. O.A. contributed to protocol design and data collection, performed the acoustic

measurement and technical validation, produced the released data tables and measurement code, and drafted the manuscript. Both authors reviewed and approved the final manuscript.

Competing Interests

Under the contract between the parties, TÜBİTAK UZAY and the Turkish Space Agency could derive financial benefit from sales of the software used in this investigation, as disclosed in the participant consent form. Conflict of interest disclosures were filed with the NASA Institutional Review Board for all investigators. The authors declare no other competing interests.

Funding

This work was supported by the Scientific and Technological Research Council of Türkiye (TÜBİTAK) and the Turkish Space Agency (TUA) through the Turkish Astronaut and Scientific Mission (TABM) Project (1007-KAMAG-121L002; Agreement No. TABM-HZT-23-28). The funders had no role in study design, data collection, analysis, or the decision to publish.

Ethics

The study was approved by the NASA Institutional Review Board (FWA No. 00019876; eIRB No. STUDY00000621, expedited review category 4) and conducted in accordance with the Declaration of Helsinki, under protocols with the Turkish Space Agency, NASA and Axiom Space. Written informed consent was obtained from all participants. The consent provides for presentation and publication of results in journals and at scientific conferences, without information that could identify participants. It also specifies that investigators do not return individual research results to participants. Accordingly, this Data Descriptor reports no individual-level clinical interpretation.

Acknowledgements

The authors thank the crew for their participation, and the Turkish Space Agency, NASA and Axiom Space for mission and protocol support.