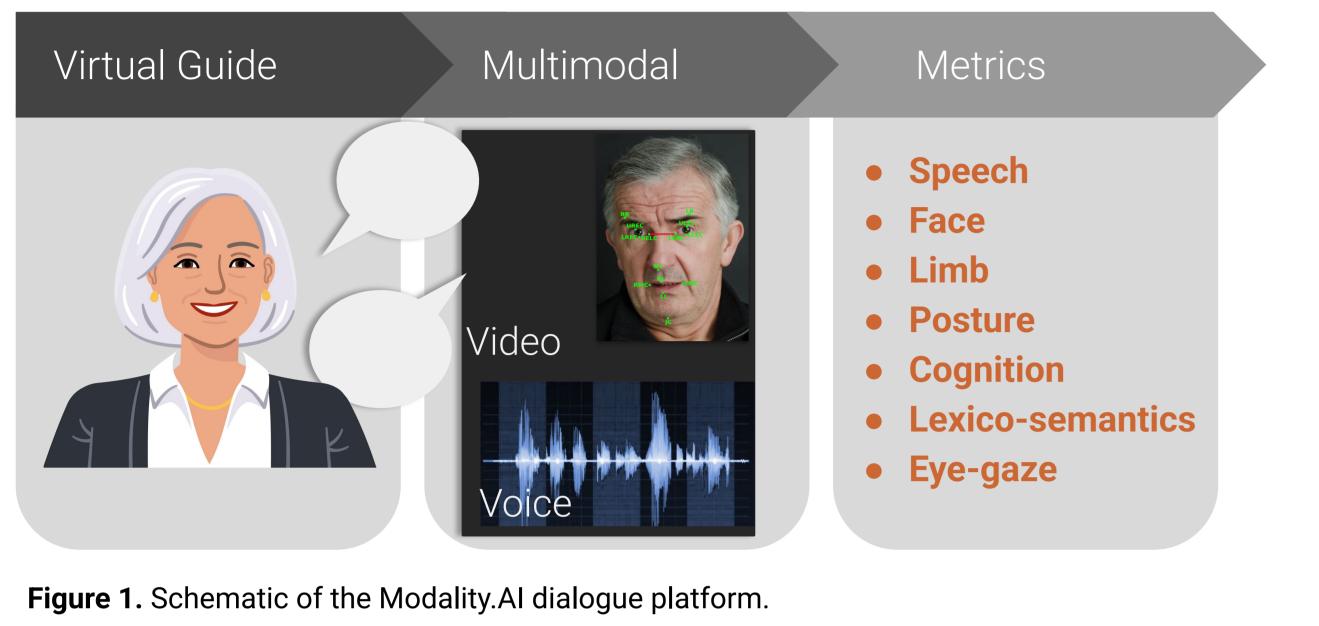




Introduction



- As part of the Target ALS Global Natural History study, the Modality.AI **platform** is used to collect digital speech, language, respiratory, facial, and limb motor function measures
- Data is collected **remotely** (at home) using a **web-based multimodal** dialog system for self-guided clinical assessments
- Novel assessment about **activities of daily living** (ADL) introduced in this multi-site, longitudinal study to assess limb motor function

Objective: Evaluate **feasibility** and **clinical utility** of **digital multimodal biomarkers** compared to traditional clinical endpoints, such as the ALSFRS-R

Data

- Tina, a **virtual guide**, walks participants through a structured set of speaking exercises and other assessment tasks
- Standard speech tasks include **sustained vowel phonation**, diadochokinesis (DDK), picture description, and reading tasks
- Optional **user experience questionnaire** is presented after the conversation with Tina
- Data collected to date: 87 recording sessions from **19 participants** within Target ALS Global Natural History study, and 58 sessions from 57 healthy controls through crowdsourcing platform Prolific

	# Participants	# Sessions	Mean Age (SD)	ALSFRS
pALS	14 (7 female)	75	62.9 (8.5)	
HC	62 (27 female)	70	52.3 (13.6)	
ALL	76 (34 female)	145	59.7 (11.1)	

Table 1. Participant statistics. *ALSFRS-R at participants' baseline session.

Toward Comprehensive Multimodal Dialog Assessment of Speech, **Respiratory, Orofacial and Limb Motor Function in ALS Patients**

Michael Neumann¹, Cathy Zhang¹, Oliver Roesler¹, Hardik Kothare¹, Jackson Liscombe¹, Meredith Bartlett¹, Jingqi Zhu², James Berry², Frances Aponte³, Brenda Deliz³, Valerie Wojna³, Manuela Quiroga Carillo⁴, Martha Peña⁴, Laura Dugom⁵, Robert Bowser⁶, Amy Easton⁵, Vikram Ramanarayanan¹

¹ Modality.AI, San Francisco, USA; ² Neurology, Massachusetts General Hospital, Boston, MA, USA; ³ University of Puerto Rico, San Juan, Puerto Rico ⁴ Instituto Roosevelt, Bogotá, Colombia; ⁵ Target ALS, New York, New York, USA; ⁶ Department of Translational Neuroscience, Barrow Neurological Institute, Phoenix, AZ, USA vikram.ramanarayanan@modality.ai

Takeaways

- 1. Multimodal biomarkers distinguish pALS from healthy controls.
- 2. Facial, speech, and linguistic biomarkers correlate with speech, bulbar and total ALSFRS-R scores, consistent with prior results on other datasets.
- 3. Novel motor measures computed from ADL tasks correlate with ALSFRS-R fine motor sub-scores

Feature Extraction

- The Modality platform automatically extracts speech, visual, and linguistic features in near real-time [1].
- Speech features are extracted with Praat and Montreal Forced Aligner; Facial features are based on MediaPipe Face Mesh landmarks; Linguistic features are computed using SpaCy based on automatic transcriptions (Whisper tiny).

• ADL tasks

- Participants are asked to mimic **brushing their teeth**, brushing their hair, and washing their face (without holding any objects).
- **ADL features** are computed based on the geometric centroid of hand landmarks extracted using MediaPipe Hands.

Ene Voi Fre Mo me Mo

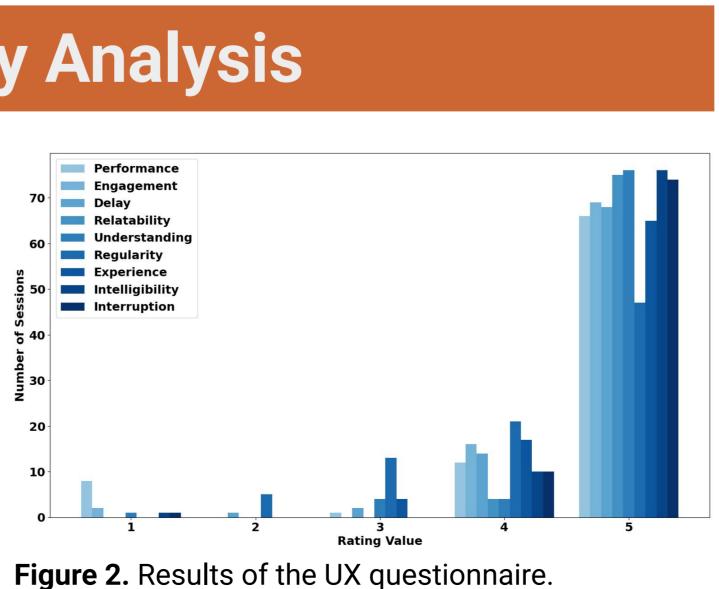
Eye

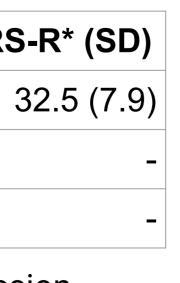
Lex sen

IOP Mo

Feasibility Analysis

- Participants were asked to rate different aspects of the **interaction** on a 5-point Likert scale.
- The majority of participants of the Natural History study rated most aspects as "Very **satisfactory**" (5).





omain	Features
nergy	shimmer (%), intensity (dB), signal-to-noise ratio (dB)
ming	speaking and articulation du- ration (sec.), articulation and speaking rate (WPM), percent pause time (PPT, %), canonical timing agreement (CTA, %)
oice quality	cepstral peak prominence (CPP, dB), harmonics-to-noise ratio (HNR, dB)
equency	mean, max., min. fundamen- tal frequency F0 (Hz), first three formants F1, F2, F3 (Hz), slope of 2nd formant (Hz/sec.), jitter (%)
outh	lip aperture/opening, lip width, mouth surface area,
easurements	mean symmetry ratio between left and right half of the mouth
ovement	velocity, acceleration, jerk, and speed of lower lip and jaw cen- ter
ves	number of eye blinks per sec., eye opening, vertical displace- ment of eyebrows
exico- mantic	word count, percentage of con- tent words, noun rate, verb rate, pronoun rate, noun-to- verb ratio, noun-to-pronoun ra- tio, closed class word ratio, idea density
ovement	cumulative distance, velocity, acceleration, and jerk of the dominant hand

Table 1. Overview of the extracted features.

- Non-parametric Kruskal-Wallis tests with **Benjamini-Hochberg** correction (Q = 0.01) were performed for each individuation feature to determine which them show a statistically significant difference betwee cohorts.
- Speech features showed the strongest signal for the Reading passage and DDK task.
- For the **ADL Face Washin**g task the cumulative distanc showed a significant difference between cohorts.
- Spearman correlations were computed between all features and related self-reported ALSFRS-R outcomes.
- Several facial features show moderate correlations with ALSFRS-R Bulbar and total score.
- A number of **speech feature** showed moderate correlation with the **ALSFRS-R speech** score and partially also with the ALSFRS-R Bulbar score.
- For the **picture description** task several linguistic featu showed moderate correlation with the **ALSFRS-R speech** total score.
- All ADL features showed moderate correlations with ALSFRS-R 6 score (dressing/hygiene).





Clinical Validation

	Reading passage - speaking duration -	
	DDK - HNR	
	Reading passage - CPP	
	Reading passage - max. mouth surface area	
	Reading passage - max. lip aperture	
	Reading passage - HNR	
	Picture description - max. mouth surface area	
lal	Picture description - max. lip aperture	
	Picture description - HNR	
of	DDK - cTV	
	Reading passage - mean F0	
	Reading passage - avg. LL speed -	
on	ADL Face Washing - Dominant Hand Distance	
en	Picture description - mean F0 -	
	Reading passage - max. LL speed -	
	Picture description - avg. LL speed -	
е	Reading passage - avg. JC speed	
	Reading passage - avg. lip aperture	
	Picture description - avg. lip aperture -	
	Picture description - avg. mouth surface area	
	Reading passage - avg. mouth surface area -	
	DDK - max. lip aperture -	
	Picture description - max. LL speed -	
	Picture description - avg. JC speed	
	DDK - max. mouth surface area	
	Picture description - CPP	
, E	Reading passage - max. lip width	
	Picture description - max. lip width	
	DDK - avg. eye opening	
5.	Picture description - avg. eye opening	
	Reading passage - avg. eye opening -	
е	DDK - syllable rate -	
_	Reading passage - CTA -	
	Reading passage - speaking rate -	

Figure 3. Effect sizes of statistically significant speech, facial, and ADL metrics. Positive effect sizes mean larger values for pALS.

	DDK - avg. JC speed -	0.33*	0.43*	0.12	- 1.00
haw	DDK - avg. LL speed -	0.24*	0.39*	0.43*	0.75
wed the	DDK - avg. mouth symmetry ratio -	-0.18	-0.07	-0.47*	- 0.75
the	DDK - max. LL speed -	0.20	0.25*	0.53*	- 0.50
	DDK - max. lip aperture -	0.36*	0.51*	0.43*	- 0.50
	DDK - max. mouth surface area -	0.37*	0.53*	0.41*	- 0.25
	DDK - syllable rate -	0.36*	0.25*	0.53*	- 0.25
	Picture description - HNR -	-0.51*	-0.42*	-0.26*	- 0.00
	Reading passage - CPP -	-0.44*	-0.34*	-0.31*	0.00
es	Reading passage - speaking duration -	-0.51*	-0.45*	-0.19	0.25
	Reading passage - speaking rate -	0.50*	0.43*	0.16	-0.25
ons	Picture description - gerundRate -	0.21	0.17	0.46*	0.50
	Picture description - inflectedverb rate -	0.18	0.13	0.43*	-0.50
	Picture description - lightverb rate -	0.20	0.16	0.41*	0.75
h	Picture description - ngraph_PE -	0.41*	0.25*	0.04	-0.75
11	Picture description - sgraph_PE -	0.44*	0.26*	0.07	1.00
2		ALSFRS-R 1 (Speech)	ALSFRS-R Bulbar	ALSFRS-R Total	-1.00
•	Fi anna A Operalation		arfacial	a maaala a	

Figure 4. Correlation results for facial, speech, and linguistic metrics. * indicates p-values below 0.05.

ures					- 1.00
ions	ADL Face Washing Dominant Hand Acceleration	0.31*	0.41*	0.29*	- 0.75
or	ADL Face Washing Dominant Hand Distance	0.39*	0.48*	0.46*	- 0.50 - 0.25
41a a	ADL Face Washing Dominant Hand Jerk	0.29*	0.40*	0.26	- 0.00
n the	ADL Face Washing Dominant Hand Velocity	0.32*	0.44*	0.33*	0.50 0.75
		ALSFRS-R 5	ALSFRS-R 6	ALSFRS-R 7	1.00

ALSERS-R 5 ALSERS-R 6 ALSERS-R 7 (Cut Food) (Dress/Hyg) (Turn Bed) Figure 5. Correlation results for ADL metrics * indicates p-values below 0.05.