Assessment of atypical speech in Multiple Sclerosis via a multimodal dialogue platform: An exploratory study

Hardik Kothare[1], Michael Neumann[1], Jackson Liscombe[1], Oliver Roesler[1], Doug Habberstad[1], William Burke[1], Andrew Cornish[1], Lakshmi Arbatti[1,2], Abhishek Hosamath[1,2], David Fox[1], David Pautler[1], David Suendermann-Oeft[1], Ira Shoulson[1,2,3], Vikram Ramanarayanan[1,4]
[1]Modality.AI, Inc., [2]Grey Matter Technologies, Inc., [3]University of Rochester, [4]University of California, San Francisco

Introduction

- The prevalence of **dysarthria in Multiple Sclerosis** (MS) is around 45%[1] with most people manifesting mild severity[2].

- Impaired speech production in people with MS (PwMS) **impacts quality of life**[3], highlighting the need to define **speech-related biomarkers for remote patient monitoring**, **tracking disease progression** and the **outcomes of therapeutic interventions**.

- **Exploratory study** investigating the feasibility of a **multimodal dialogue platform** with **real-time extraction of speech acoustic and facial kinematic metrics** in assessing impaired speech motor control in MS.

Methods and Materials

- **9 PwMS and 9 age-matched controls (all female**, Table 1) completed an interactive session in December 2021 and January 2022 using a cloud-based multimodal dialogue platform (Illustration in Figure 1).

- Participants were guided through a **battery of tasks** eliciting speech and facial behaviours: sustained vowel phonation, counting up numbers in a single breath, repeating consonant-vowel-consonant (CVC) words, alternating-motion rate diadochokinesis, reading sentences and passages, picture description, spontaneous speech on a topic of their choice.

- **Survey instruments** at the end of the interactive session: the short form of the Communicative Participation Item Bank (CPIB-S), the Schwab and England Activities of Daily Living scale and the Patient Report of Problems (PROP™).

- Speech acoustic and facial kinematic metrics were automatically extracted (Table 2). Facial metrics were normalised for each participant by the inter-caruncular distance between the eyes. **Non-parametric Kruskal-Wallis tests** were performed to investigate differences between PwMS and controls.

	Number of participants	Mean age ± standard deviation (years)	Median Schwab and England score (Q1-Q3)	Median CPIB-S score (Q1-Q3)
PwMS	9 female (7 RRMS, 2 SPMS)*	40.22 ± 8.44	90 (70 - 90)	5 (4 - 9)
Controls	9 female	40.11 ± 8.25	100 (80 - 100)	1 (0 - 10)

Table 1: Demographics
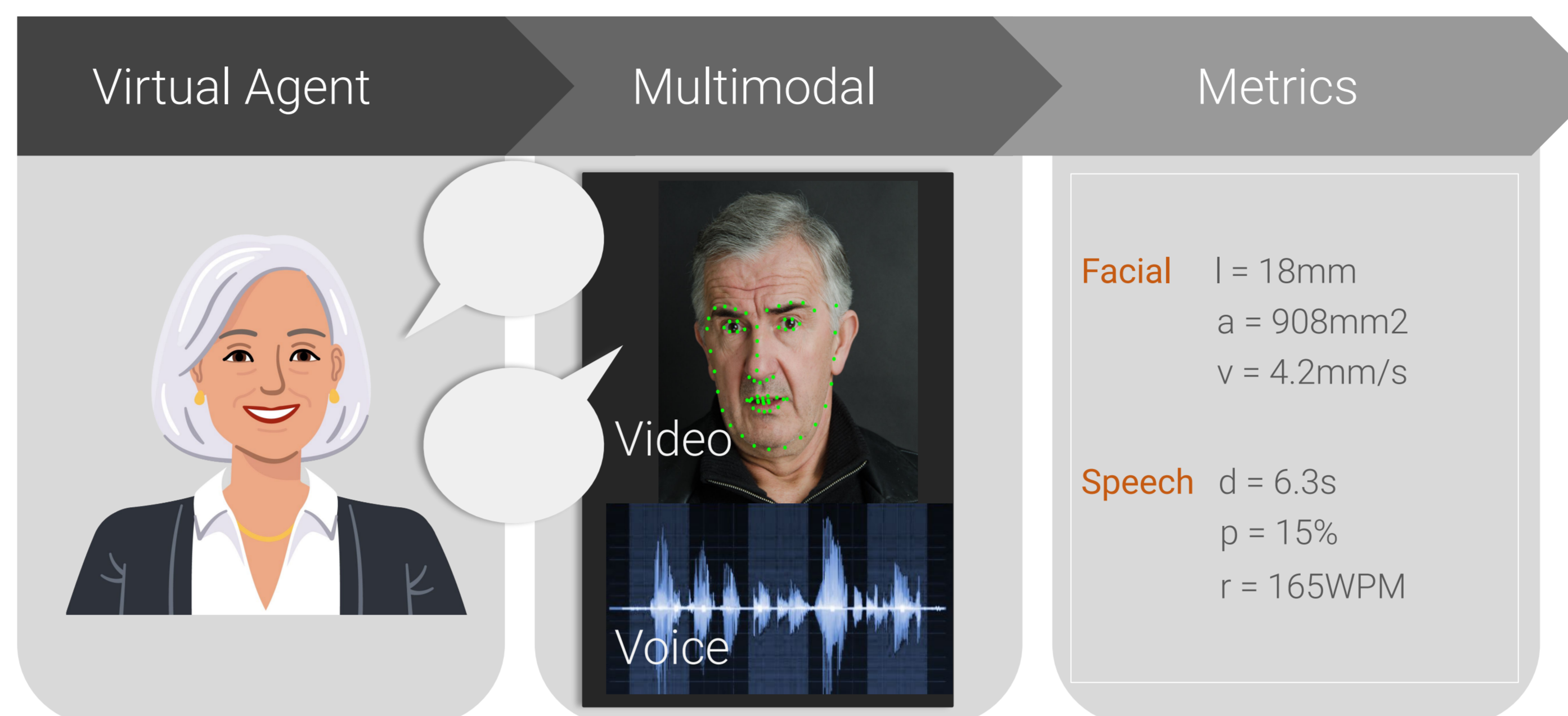* RRMS = Relapsing-Remitting MS, SPMS = Secondary Progressive MS

Figure 1. Modality.AI dialogue platform.

Acoustic measures	- **Fundamental Frequency (F0):** Minimum value (Hz) and timepoint (s), Maximum value (Hz) and timepoint (s), Mean (Hz), Standard Deviation (Hz) - **Formant Frequency Values:** F1, F2, F3 (Hz) and F2 slope (Hz/s) - **Cepstral Peak Prominence** (CPP in dB) - **Harmonics-to-Noise Ratio** (HNR in dB) - **Articulation duration** (in s, excluding pauses) and **speaking duration** (in s, including pauses) - **Articulation rate** and **speaking rate** (words per minute) - **Percent pause duration** (%) - **Signal-to-noise ratio** (SNR in dB) - **Articulation intensity** (dB) - **Jitter** and **shimmer** (%)
Visual measures	velocity, acceleration, and jerk of lower lip and jaw center, lip aperture, lip width, eye opening, vertical eyebrow displacement, eye blinks, area of the mouth, symmetry ratio of the mouth area

Table 2. Automatically extracted acoustic & visual measures.

Results and Discussion

- A variety of metrics showed statistically significant differences between PwMS and controls (Figure 2) at an **alpha threshold of 0.01** and were **controlled for false discovery rate**.

- PwMS showed **greater values of higher-order derivatives of the vertical movement of the jaw** (acceleration and jerk) during the production of /ɛ/ and /i/ CVC words, indicating **lack of smoothness in movement**.

- PwMS exhibited **shorter articulatory duration during spontaneous speech** production accompanied by a **larger percentage of pause duration**.

- **Wider mouth opening** in PwMS during **sustained phonation of /ɑ/**.

- **Lower cepstral peak prominence (CPP)** in PwMS during sustained phonation of /ɑ/, indicating a **relative degradation in voice quality**.
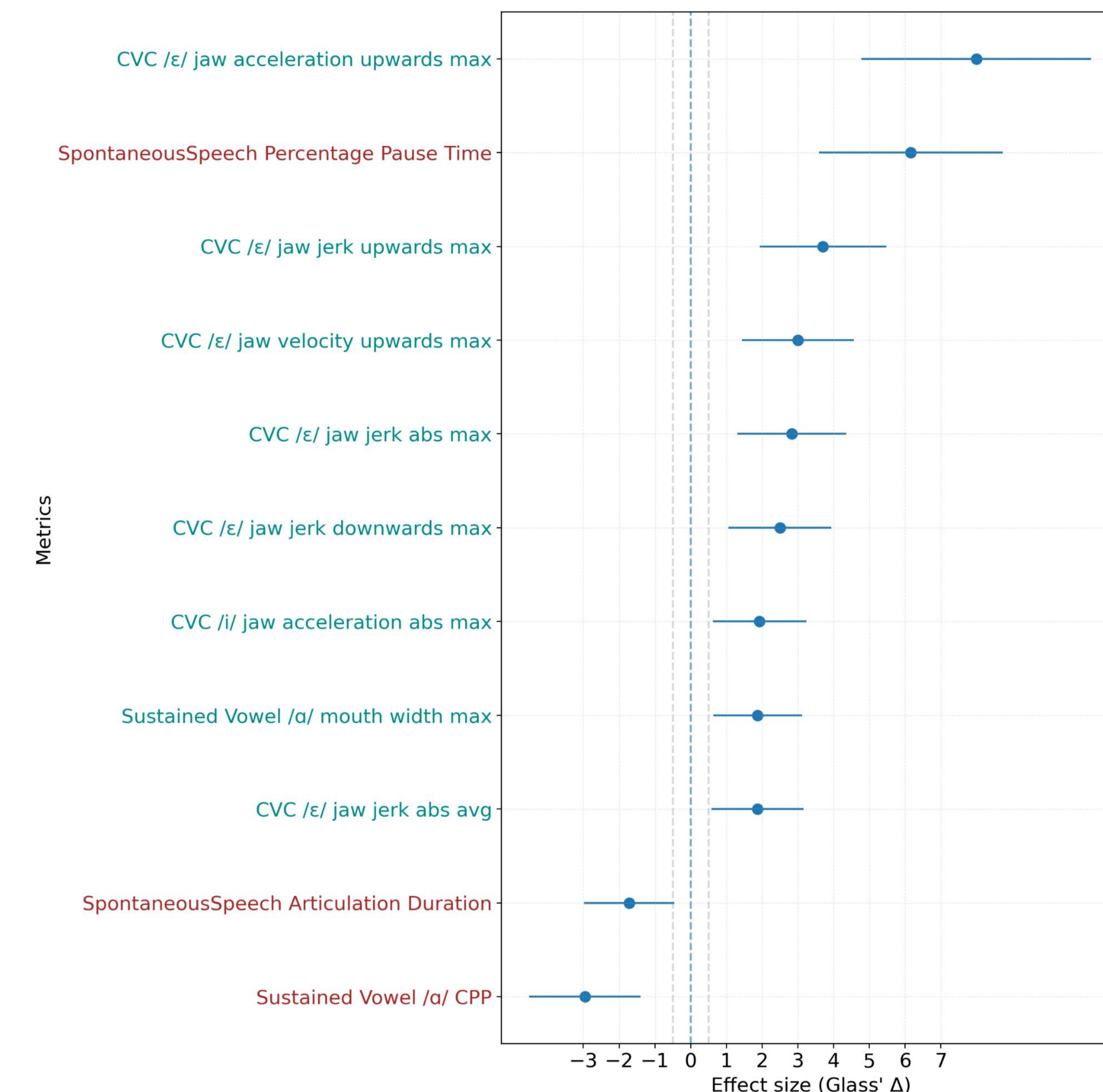

Figure 2. Effect sizes of **acoustic** and **facial** metrics that show statistically significant differences between PwMS and controls at an alpha threshold of 0.01.

Conclusions and Limitations

- These findings support the **feasibility of assessing and monitoring objective measures of atypical speech production in MS** through the use of a novel multimodal conversational technology.

- The sample size in this exploratory study is very small and future studies with larger cohorts will be needed to confirm the findings.

References

1. Noffs, Gustavo, et al. "What speech can tell us: A systematic review of dysarthria characteristics in Multiple Sclerosis." Autoimmunity reviews 17, no. 12 (2018): 1202-1209.
2. Fitz Gerald, Fiona J., et al. "Multiple sclerosis: Associated speech and language disorders." Australian Journal of Human Communication Disorders 15, no. 2 (1987): 15-35.
3. Piacentini, Valentina, et al. "Relationship between quality of life and dysarthria in patients with multiple sclerosis." Archives of physical medicine and rehabilitation 95, no. 11 (2014): 2047-2054.
4. Ramanarayanan, V., et al., 2020. Toward Remote Patient Monitoring of Speech, Video, Cognitive and Respiratory Biomarkers Using Multimodal Dialog Technology. In INTERSPEECH (pp. 492-493).